CUK-H05880-7-261212

CALC
CEN
LIBR
UNIVE

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

370.15
B

FEBRUARY 1966

VOL. XXXVI, PART 1

7

Price One Pound

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
L. B. BIRCH

Associate Editors:
P. E. VERNON and DORIS M. LEE

Advised by

SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL
SIR FRED J. SCHONELL
S. WISEMAN

Vol. XXXVI, Part 1

The British Journal of Educational Psychology is issued by the British Psychological Society and the Association of Teachers in Colleges and Departments of Education, once a term, in February, June and November. The subscription is £2 a year post free payable in advance, or £1 each number.

Subscriptions and Orders should be sent to "The Manager, B.J.E.P. Department, Methuen and Co., Ltd., 11, New Fetter Lane, London, E.C 4. Members of the British Psychological Society and members of the A.T.C.D.E. receive the *Journal* on special terms through the Societies.

Papers for publication should be sent to L. B. Birch, The Institute of Education, The University, Sheffield, 10. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to L. B. Birch, at the above address. *Advertisements* and *Subscriptions* for this *Journal* should be sent to Manager, B.J.E.P. Department, Methuen and Co., 11, New Fetter Lane, London, E.C.4.

Published by
METHUEN AND CO., LTD., 11, New Fetter Lane, London, E.C.4.,
for the British Journal of Educational Psychology, Ltd.

Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

261212

VOLUME XXXVI FEBRUARY, 1966 PART 1

A STUDY OF THE EFFECTS OF INSTRUCTIONS IN A MULTIPLE-CHOICE MATHEMATICS TEST

By P. H. Taylor
(*University of Leicester*)

Summary. A moderately speeded, moderately difficult multiple-choice mathematics test was set under three different instructions to three random samples of 15-year-old pupils of average and above average ability. The aim of the instructions was to dispose the pupil so as to inhibit his propensity to guess (A), to encourage it (C), and to be neutral in respect to guessing (B).

Analysis of variance of both raw and corrected scores showed there to be no significant effect attributable to instructions either in the whole test or in the earlier or later part of the test The instructions affected the pupils' item response pattern. There were fewer ' omits and unfinished ' items than would be expected by chance arising from the inhibiting instructions (A), more from the instruction encouraging guessing (C) and the level expected by chance from the neutral instructions (B). The concurrent validity was not significantly affected by the instructions.

7

I.—Introduction.

Development work related to the Certificate of Secondary Education examination indicated that one factor in the response of teachers to the use of objective type tests in the examination was a fear that pupils would obtain high scores by guessing alone (H.M.S.O. 1964). Evidence related to the propensity of pupils to guess (H.M.S.O. 1965) in a multiple-choice mathematics test suggested that there was no ground for supposing that high scores were obtained by guessing alone or that pupils guessed at random. However, the evidence though in keeping with other findings had nothing to say on the effect of instructions to pupils on their propensity to guess and on the distribution of scores arising from tests having different instructions to pupils.

There have been a number of studies in the U.S.A. of the effects of instructions to pupils set multiple-choice objective type tests. This work has, in most cases, been related to the problem of guessing. Gulliksen (1950), Fredericksen (1952), Traxler (1956), and Lord (1963) suggest that as pupils' responses to items are non-random, even when they guess, no correction for random guessing is necessary as long as the test is of an appropriate level of difficulty, is not markedly speeded and the instructions do not direct the pupil to guess wildly. Lord (1964) suggested that instructions aimed to have the pupil attempt every item would give rise to a spurious increase in test validity arising from random error. However, no study had investigated the effects of the instructions given to examinees on the means and variances of multiple-choice test scores, the effect on the item-response behaviour of pupils and on the validity of the test at one and the same time. The development work in the field of the Certificate of Secondary Education examination provided such an opportunity.

II.—Plan of the study.

Three sets of instructions were drawn up each to represent one of three possible modes of response to multiple choice test items: a discouragement to guessing, a neutral response to guessing and an encouragement to guessing.

Each of the instructions was preceded by the statement 'Work through the questions in the order in which they are presented in the test paper'; and followed by the statement: 'When you have finished the last question, please hand in your paper.'

These instructions, A, B, and C, were printed on both the outside and the inside of the front cover of the test booklet and were as follows:

A. *Do not* record an answer until you have studied the question carefully and *do not record an answer unless you are certain it is correct.*

B. *Your aim should be to do as well as you can.*

C. It is important that you get as many right as possible. In order to do this you may need to guess what the answer is.
Don't be afraid to guess and to have a go at every question.

The test was a moderately speeded (1 hour), moderately difficult (Md.F=48·5 per cent. ranging from items with F=8·5 per cent. F=85·4 per cent and moderately discriminating (Md. D=0·41) test of forty-five items, and was composed of nine mathematical situations of five facets, of which one-third were in each of the areas geometry, arithmetic and algebra. The items were, in fact, selected from a longer multiple-choice, multi facet test which had been subject of an earlier study (H.M.S.O. In press). A ten-minute Practice Test designed to familiarise the pupils with the form of the test was also constructed and administered.

Teachers in twenty-three secondary schools in the south west of England co-operated in the study, and the total sample of pupils tested was 830. *Table* 1 gives the distribution of the sample by type of school. The pupils tested were in the ability range of 40th to 99th percentile, with the majority being of average ability or a little above, and were in their fifth year at school.

TABLE 1

DISTRIBUTION OF SCHOOLS AND PUPILS IN THE SAMPLE N=830.

Type of School	Sex of School		
	Mixed	Boys	Girls
Grammar	—	—	1 (82)
Comprehensive	3 (241)	1 (63)	1 (21)
Bi-lateral	1 (30)	—	—
Modern	11 (259)	3 (116)	2 (18)
Total	15 (530)	4 (179)	4 (121)

The school was taken as the basic sampling unit. Within each school pupils to be tested were allocated at random a test having one of three instructions A, B or C. This gave 280 pupils being set the test with instruction A, 280 instruction B and 270 with instruction C. Teachers were given standard instructions for the conduct of the testing prior to which they were asked to submit a rank order of merit of pupils within the teaching groups and within schools.

III.—STATEMENT OF RESULTS.

(*a*) *Raw Scores.* The forty-five items of the test were scored 0 or 1 and gave rise to a distribution of scores within schools, sexes and instructions. The raw scores were subjected to an analysis of variance (Edwards 1963) for mixed and single-sexed schools separately. The results of these analyses are given in Table 2 below.

TABLE 2

ANALYSES OF VARIANCE OF RAW SCORES.

Type of School	Source of Variance	d.f.	Sums of Sq.	Mean Sq.	F.	Sign.
Mixed (N=492)	Between Instructions	2	133·5	66·7	1·2	N.S.
	Between Sexes	1	348·7	348·7	6·1	05
	Between Schools	12	2811·9	234·3	4·1	01
	Instructions and Sexes	2	106·0	53·0	0·9	N.S.
	Sexes and Schools	12	953·1	79·4	1·4	N.S.
	Schools and Instructions	24	702·2	29·3	0·5	N.S.
	Sexes and Schools and Instructions	24	696·5	29·0	0·5	N.S.
	Residual (within groups)	414	23615·2	57.0	—	—
	Total	491	29367·1	—	—	—
Boys (N=179) (N=179)	Between Instructions	2	12·7	6·3	—	N.S.
	Between Schools	3	601·7	200·6	2·9	05
	Instructions and Schools	6	120·1	20·0	—	N.S.
	Residual (within groups)	167	11631·9	69·6	—	—
	Total	178	12366·4	—	—	—
Girls (N=121)	Between Instruction	2	136·9	68·5	1·1	N.S.
	Between Schools	3	2281·6	760·5	12·4	01
	Instructions and Schools ...	6	224·4	37·4	—	N.S.
	Residual (within groups)	109	6693·0	61·4	—	—
	Total	120	9336·7	—	—	—

(*b*) *Scores Corrected for Guessing.* The raw scores were corrected for guessing (Lindquist 1956) and the results subjected to analysis, again separately for mixed and single sexed schools. The results are given in Table 3.

(*c*) *Scores for First and Last Thirds of the Test.* The analyses of variance of both raw and corrected scores for the total test indicated that the instructions were not significant at the 5 per cent. level, so similar analyses were carried out on the first and last thirds of the test. Again there were no significant effects due to instruction. The results are, therefore, not reported in detail.

(*d*) *Items Omitted and Unfinished.* It was clear at the outset that although the instructions might not affect the means of the scores, they might, nevertheless, affect other relevant aspects of a candidate's behaviour, i.e., his propensity to omit items and to finish the test. The table below indicates that the instructions had a different effect on the numbers of items omitted and unfinished. In every case chi square was significant at the 1 per cent. level. The most reliable data is probably the last where omitted and unfinished items are taken together. This is so because it is not always clear whether an item was omitted or not reached.

TABLE 3

ANALYSES OF VARIANCE OF CORRECTED SCORES.

Type of School	Source of Variance	d.f.	Sums Sq.	Mean Sq	F.	Sign.
Mixed (N=492)	Between Instructions	2	90·2	45·1	0·5	N.S.
	Between Sexes	1	397·4	397·4	4·4	05
	Between Schools	12	5140·4	428·4	4·7	01
	Instructions and Sexes	2	78·8	39·4	0·4	N.S.
	Sexes and Schools	12	2009·4	167·4	1·8	N.S.
	Schools and Instructions	24	1125·5	46·9	·5	N.S.
	Sexes and Schools and Instructions	24	959·6	40·0	·4	N.S
	Residual (within groups)	414	37664·5	90·1	—	—
	Total	491	47465·8	—	—	—
Boys (N=178)	Between Instructions	2	127·1	63·5	—	N.S.
	Between Schools	3	1340·4	446·8	3·9	01
	Instructions and Schools	6	83·9	14·0	—	N.S.
	Residual	167	19016·4	113·9	—	N.S.
	Total	178	20567·8	—	—	—
Girls (N=121)	Between Instructions	2	229·1	114·5	1·2	N.S.
	Between Schools	3	3628·4	1209·5	12·6	01
	Instructions and Schools	6	450·2	75·0	—	N.S
	Residual	109	10467·9	96·0	—	—
	Total	120	14775·6	—	—	—

TABLE 4

CHI SQUARE FOR ITEMS OMITTED; UNFINISHED AND ITEMS OMITTED AND UNFINISHED (BOYS AND GIRLS).

Instructions		A	B	C	Total	
Omitted	O E	1329 947·3	1010 947·3	469 913·4	2808 2808	$X^2=374{\cdot}2$, $P<{\cdot}01$
Unfinished	O E	610 414·6	314 41·46	305 399·8	1229 1229	$X^2=139{\cdot}0$, $P<{\cdot}01$
Omitted and Unfinished	O E	1939 1361·9	1324 1361·9	774 1313·2	4037 4037	$X^2=467{\cdot}0$, $P<{\cdot}01$

(*e*) *Concurrent Validity*. Using teachers' rank orders of merit it was possible by converting the order of merit into scores (Garrett 1958) to correlate within school rankings with within school scores, both raw and corrected, and using the Z-score transformation to arrive at indices of concurrent validity for each of the three instructions. These were :

Instruction.

	A	B	C
(*a*) Raw Scores	0·69	0·75	0·77
(*b*) Corrected Scores	0·70	0·76	0·77

(The differences within (*a*) and (*b*), between instructions, and within instructions A, B, and C between forms of scoring are nowhere significant at the 5 per cent. level).

IV.—DISCUSSION AND CONCLUSION.

It would seem reasonable to suggest that there is no effect on the means and variances of the scores, whether raw or corrected, arising from the differences in instruction. This would appear to be equally as true for both the first and last parts of test as it is for the test as a whole. These findings are in keeping with other similar work on the effect of instructions (Frederiksen 1952 and Lord 1963).

On the other hand, as the data on items omitted and unfinished suggested, the instructions do have a differential effect on the item response behaviour of candidates. Instruction A, the cautionary and inhibiting instruction, produces significantly more ' omitted and unfinished ' items than one would expect by chance, whereas instruction C, the encouraging and uninhibiting instruction, produces significantly fewer. The neutral instruction B produces the response one would expect by chance. The pattern is somewhat different when one considers the data on items ' omitted ' and ' items unfinished ' but, for reasons stated earlier, this may be less reliable data. Thus, one effect of the different instructions is to change the response behaviour of the pupils being tested. Pupils finish and/or complete more or fewer items ; or about the same as one would expect by chance, depending upon the instructions. The instructions do not, however, affect the means and variations of their scores no matter what the pupil's sex or the type of school which he attends.

If it is reasonable to suppose that the instructions used in this study affected the propensity of pupils to guess, then the results, especially where a correction of the raw scores for random guessing was applied, suggest that the pupils studied faced with a multiple-choice test do not guess at random, and conversely, that an instruction inhibiting their propensity to guess does not affect their scores relative to the scores of pupils acting under other instructions. It would seem, therefore, that right answers, over and above right answers obtained by normal responses to items, do not accrue to pupils who are instructed to guess. Their guessing would seem to be non-random. Vernon's (1964) statement that examinees seldom guess at random would seem to be borne out.

Further, the concurrent validity of the test appears to remain unaffected by the differences in instructions. This finding would not seem to be in keeping with Lord's evidence (Lord 1964) in which he states ' Non-random forced guessing, however, may increase validity by virtue of its non-randomness.' There is, however, the point that the 5 per cent. level of significance used to test the differences between the indices of concurrent validity may be too severe. The gains and losses in concurrent validity for the two extreme instructions A and C relative to the neutral instruction B were in the same direction and of much the same order for raw and for corrected scores, as may be seen below :

	A	B	C
(*a*) Raw Scores	−0·06	0·00	+0·02
(*b*) Corrected scores	−0·06	0·00	+0·01

These suggest that Lord's warning—that " proposals requiring every examinee to answer every item should be viewed with some caution " should not be ignored despite the fact that the differences in concurrent validity found in this study were not significant at the 5 per cent level. It appears, therefore, that the balance of evidence suggests that instructions which may distort the pupil's normal or unforced response to items by requiring him " to have a go at every question " (Instruction C) should be avoided. On the other hand the chance of error affecting the validity of the test does not seem to be as high where the instructions require the pupil to exercise great caution in his response

(Instruction A). There is, however, the possibility that because he has been inhibited from answering items he would like to have tried but did not, the pupil will feel that the test has been ' unfair.'

Instructions which are motivationally positive (Instruction B) but give the pupil no guidance as to his mode of response to items do not run the risk of giving rise to random error and may not leave an unnecessary feeling of unfairness. Moreover, such instructions are congruent with the motivational set encouraged by the general run of teachers in the work the pupil does in school. Any backwash effect of such instructions is likely to re-inforce the efforts of the teacher, and, therefore, be to the good.

This study does not resolve the problem of whether to apply a correction for random guessing to test scores except in so far as it suggests that pupils in general do not guess at random. Gulliksen and others (Gulliksen 1950, Traxler 1956) do not advocate the use of a correction for random guessing unless there is evidence that the policy among pupils is to guess at random. There is evidence (H.M.S.O. 1965) that such a policy does not yet exist among pupils in this country at least as far as this kind of mathematics test is concerned.

Perhaps Lord's (Lord 1963) criteria remain the best guide: apply a correction: (1) when there is wide variability among examinees in the propensity to guess, (2) when there are fewer than five choices per item and (3) when the test is quite difficult for the group tested.

It would seem reasonable to suggest as a final and general conclusion of this study that instructions which are neutral as to the mode of item response to be adopted by the pupil but are congruent with the positive motivational set of the class-room should be used in objective tests of attainment with secondary school pupils, and that unless there is evidence to suggest that there is a marked propensity among pupils to guess, or there are fewer than five choices per item or the test is quite difficult, no correction for guessing should be applied to the scores.

ACKNOWLEDGEMENTS.—The Author wishes to thank the teachers of the area of the South West Regional C.S.E. Board who made this study possible, and Mrs. R. Palmer, M.B.E., and Miss L. Paul who so conscientiously helped with the processing of the data.

V.—REFERENCES.

EDWARDS, A. L. (1963). *Experimental Design in Psychological Research.* New York: Holt, Rinehart and Winston, vii-398.

H.M.S.O. (1964). *Examinations Bulletin No. 2. The Certificate of Secondary Education: Experimental Examination—Mathematics.* H.M.S.O., London, p iii-35.

H.M.S.O. (1965). *Examinations Bulletin No. 8: Experimental Examinations—Mathematics* 2. H.M.S.O., London (in press).

FREDERIKSEN, N. (1952). The influence of timing and instructions on co-operative reading test scores. *Educ. Psychol M'ment,* Vol. 12, No 4, pp. 598-607.

GARRETT, H. E. (1958). *Statistics in Psychology and Education.* New York: David McKay, v-478.

GULLIKSEN, H. (1950). *Theory of Mental Tests.* New York: John Wiley, p. iii-486.

LORD, F. M. (1963). *Formula scoring and validity. Educ. Psychol. M'ment,* Vol. 23, No. 4, p. 663-672.

LORD, F. M. (1964). The effect of random guessing on test validity. *Educ. Psychol. M'ment,* Vol. 24, No. 4, pp. 745-747.

VERNON, P. E (1964). *Examinations Bulletin No. 4. The Certificate of Secondary Education: An Introduction to Objective-type Examinations.* H.M.S O., London, p. iii-20.

(*Manuscript received 17th May,* 1965.)

SECONDARY EDUCATION, SOCIAL CLASS AND THE DEVELOPMENT OF ABILITIES

By W. B. DOCKRELL

SUMMARY. This paper reports an investigation into the relationship between patterns of abilities and two environmental variables, socio-economic status, and type of secondary education for 12 and 14-year-olds. It was hypothesized that education and social class will interact so that middle-class children will show no change of pattern of abilities in academic high schools, but in technical high schools will gain most rapidly in spatial and other non-verbal skills ; lower class children will gain in verbal skills in the academic schools and in non-verbal skills in the technical schools for none of the academic abilities of lower-class children are fully developed in the out-of-school environment. These results are interpreted to mean that the environmental demands in a middle-class home are such that a high level of development has been reached in all areas and that type of secondary education is, therefore, irrelevant, but the lower-class home does not demand as full a level of development, particularly in verbal areas, and the type of secondary school can, therefore, markedly influence intellectual development

I.—INTRODUCTION.

DISCUSSIONS of secondary education whether it be of the ' types of mind ' approach of the Hadow (Great Britain, Ministry of Education 1926) and Norwood (Great Britain, Ministry of Education, 1943) reports, or the levels of ability approach Spens (Great Britain, Ministry of Education, 1938), Crowther, (Great Britain, Ministry of Education, 1959) and Newsom (Great Britain Ministry of Education, 1963) reports and Conant (1959) and Rickover 'U.S. Congress, 1963) usually assume the ' empty jug ' concept of intelligence. The jug may be constitutional or environmental, but whichever it is, education fills and does not shape it. There has been some recognition that length of education may be important (Great Britain, 1963) and there is evidence to support this view (Vernon, 1947 ; Husen and Svennson, 1960). But there has been little consideration of the extent to which a particular type of education may influence the development of particular abilities.

Studies usually cited as most relevant to the extent of environmental influences are investigations of identical twins raised separately and unrelated children reared together (Warburton, 1958). These data, however, are by no means unambiguous. Though height shows no correlation with the environmental variables, measured intelligence can be interpreted as showing nearly 30 per cent. of the variance contributed by the environment in the case of the identical twins reared separately.

There is no agreement on which aspects of the environment are important. Hebb and his associates, supported by a considerable body of research evidence, stress the " lasting and generalized effects of early learning " (Hebb 1949, p. 299). Attempts to relate level of intelligence to specific aspects of the cultural background, have in general been unsuccessful (Osborn 1947 ; Fraser 1959) though the role of general culture factors (Deutch 1960) and parental attitudes (Kent and Davies 1957 ; Jackson and Marsden 1962) has been substantiated. As with the importance of environment in general, the either or position on the importance of early learning does not seem tenable.

Hebb (1949) does not believe, that the formal school curriculum develops intelligence. Burt (1959), however, argues that the school has a stimulating effect on lower class children. There is evidence that the usual elementary school curriculum does not compensate for the effects of early cultural deprivation (Dockrell 1958) though specially designed programmes may (Dawe 1942). The importance of education in adolescence, however, is borne out by a number of studies (Vernon 1947 ; Clarke and Clarke 1958 ; and Husen and Svennson 1960). Husen further showed that the effects of educational experience in later childhood and adolescence were specific rather than general and Fillela (1960) interpreted his findings in the same way.

Though theory is contradictory and the research evidence ambiguous, Ferguson's formulation (1954, 1956) offers a framework for further study. For Ferguson " abilities are overleared acquisitions " (1954, p. 98). What abilities will be developed is a function of the demands of a particular culture. " Presumably children reared in different environments, which demand different types of learning at different ages develop different patterns of abilities" (p. 98). When this theory is applied to the impact of secondary education on children of different social class backgrounds it suggests that middle-class children who are presumably at or close to their limit in development of verbal abilities will not be influenced by formal schooling in this area but in areas where their abilities are not fully developed educational influences will be important. Lower class children on the other hand, whose verbal skills will not have been fully developed, will react strongly to an educational environment that demands these abilities. If, as seems likely, the out-of-school environment has not stimulated to the full any of the abilities of lower-class children educational factors should be important even in the development of non-verbal abilities.

II.—THE INVESTIGATION.

The study was conducted in a single community chosen to provide as wide a range of social class groups as is available in the maintained schools. The distribution of population by census classifications closely approximated the national figure. Occupation of the breadwinner was chosen as the basis for assessment of social class but the scale devised by Moser and Hall (1954), and not the census groupings, were used in the study. Their groups A, B, and C were combined to provide a middle-class group and groups D and E were combined to provide a lower-class group. The middle-class group included professional, managerial and clerical workers and the lower-class skilled and unskilled workers.

The second criterion for choice of the community was educational. The secondary schools were clearly differentiated and well established. The proportion of private school places available in the community was below the national average, less than 1 per cent., though this figure was no doubt inflated by a number of children attending boarding schools and others attending independent or direct grant schools in the large urban centre some thirty miles away. For the study reported here, the schools chosen were the two grammar schools, boys and girls ; the secondary technical school ; and the two modern schools serving the predominantly middle-class neighbourhood and a working-class area. In each school all pupils in the appropriate age categories were tested. This involved testing some individuals from classes, not tested as a whole and discarding others from the classes tested as a group. The mean age of the groups was 146·4 months and 170·9 months, and the range was 12 months in each case.

The test battery consisted of Raven's Matrices, Vernon's Arithmetic-Mathematics Test, the Otis Mental Ability Test, Schonell's diagnostic English Test, Peel's Test of Practical Ability, and Thurstone's Primary Mental Ability Battery. The P.M.A. battery consists of Verbal Meaning, Space, Reasoning, Number and Word Fluency tests.

III.—RESULTS.

A—*Grammar School Pupils.*

TABLE 1

MEANS AND STANDARD DEVIATIONS OF GRAMMAR SCHOOL PUPILS DIVIDED BY SOCIAL CLASS AND BY AGE.

	Middle-Class				Lower-Class			
	Age 12		Age 14		Age 12		Age 14	
1. Raven	56·0	6·1	53·7	9·4	55·1	8·8	52 9	9·7
2. Vernon	57·8	5·9	55·4	7·7	57·5	5·4	55 7	7·8
3. Otis	58·4	5·3	56·9	5·0	57·3	7·8	56·6	5·5
4. Peel	60·5	7·2	55·2	6·5	61·3	7 2	54·4	7·6
5. Schonell	59·1	4·2	56·8	3·3	58·4	3·5	56·4	3·9
6. P M.A. V	57·9	7·0	56·4	7·9	58·4	5·7	55·6	7·5
7 S	55·2	9·5	54·3	10·3	55·4	7·8	51·2	10·0
8. R	56·9	8·5	54·1	8·4	55·4	8·4	54·5	8·6
9. N	54·9	9·5	52·2	9·0	54·5	9·0	53·7	8·6
10. W	54·0	8·8	52·6	8·0	55·4	7·6	54·7	11·4
N	96		75		52		49	

Analysis of Variance

1. Middle-Class Group :

Between ages	P	0·01
Between tests	P	0·01
Interaction	N.S.	

2 Lower-Class Group :

Between ages	P	0·01
Between tests	P	0·01
Interaction	P	0·01

The analysis of data was made by the Computing Centre of the University of Alberta on the IBM 1620 computer. Raw scores were converted to scaled scores with a mean of 50 and a standard deviation of 10 at each age level. Tests of homogeneity of variance and corrections for unequal cell frequencies were made before the analyses of variance were computed. The results for the three types of secondary schools are presented separately.

Because the scores are standard scores based on groups of slightly different composition any general trends, for example a general improvement of grammar school pupils would be concealed. However, a shift on some tests and not on others, that is a change in pattern, is shown by the inter-action term. The significant differences between ages is an artifact of the data. The significant difference between tests indicates that the grammar school children made a higher average score on some tests than on others. For the purposes of this study the important point is the inter-action term. This indicates whether there has been a greater shift on some tests than on others, whether there is a difference in pattern of abilities at different ages. The middle-class children in the grammar school show no change in pattern, the lower-class children show a significant change. Table 2 shows the nature of these changes.

TABLE 2

TESTS RANKING IN ORDER OF RELATIVE GAIN BETWEEN TWELVE AND FOURTEEN OF LOWER CLASS GRAMMAR SCHOOL CHILDREN.

Otis		
P.M.A.	N	
P.M.A.	W	
P.M.A.	R	
Vernon		
Schonell		
P.M.A.	V	a
Raven		b
P.M.A.	S	b
Peel		b

a differences significant at the 0·05 level.

b differences significant at the 0·01 level.

In the grammar school the lower class children show greatest relative improvement on Otis, Vernon and Schonell of the British tests and least on Peel. On the P.M.A. battery the greatest gains were on N, W, and R, and least on S. As predicted, the middle-class children did not change in pattern of ability in response to the demands of the grammar school curriculum. The lower-class children, however did, They improved more in the verbal academic area and less on the non-verbal and spatial tests.

B—*Technical School Pupils.*

TABLE 3

MEANS AND STANDARD DEVIATIONS OF TECHNICAL SCHOOL PUPILS DIVIDED BY AGE AND SOCIAL CLASS.

			Middle Class				Lower-Class			
			Age 12		Age 14		Age 12		Age 14	
1	Raven		51·2	6·7	52·1	6·2	52·1	6·7	54·6	5·7
2.	Vernon		54·7	6·0	54·7	6·0	56·3	6·1	54·9	6·1
3.	Otis		53·6	5·9	53·0	6·3	56·7	4·5	53·0	5·8
4.	Peel		56·3	10·2	55·8	7·7	58·7	8·3	56·0	6·3
5.	Schonell		55·9	3·5	53·7	4·0	56·0	3·7	53·3	4·6
6.	P.M.A.	V	56·4	8·4	54·2	7·9	53·6	6·4	51·6	6·8
7		S	51·6	6·8	54·4	6·3	53·5	7·7	54·0	8·6
8.		R	49·3	7·8	53·7	7·6	53·6	7·1	52·2	8·2
9		N	54·5	9·8	49·8	10·2	54·2	8·2	53·4	9·0
10.		W.	51·8	6·7	52·5	9·1	53·7	7·1	50·8	7·3
	N		15		17		49		68	

Analysis of Variance

1. Middle-Class Group

Between ages	N.S
Between Tests	N.S.
Interaction	N.S.

2 Lower-Class Group

Between ages	P	0·01
Between tests	P	0·01
Interaction	P	0·01

Again the between ages term is an artifact of the data, and the between tests term indicates that the technical school pupils differ from the total group more on some tests than on others. What is important for the study is the interaction term. The middle-class group in the technical school is so small that no firm conclusions can be drawn. Any differences in pattern of ability resulting from the technical school curriculum do not attain statistical significance in this study. The lower-class children do show a significant shift in pattern. Table 4 shows the nature of the changes.

TABLE 4

Tests Ranked In Order of Relative Gain Between Twelve andFourteen of Lower Class Technical School Children.

Raven		
P.M.A.	S	
P.M.A.	N	
Vernon		
P.M.A.	R	
P.M.A.	V	
Peel		b
Schonell		b
P.M.A.	W	b
Otis		b

a differences significant at the 0·05 level.
b differences significant at the 0·01 level.

In the technical school the lower-class children show greatest relative improvement on Raven and Vernon of the British tests and least on Schonell and Otis. On the P.M.A. battery the greatest gains were on S, N and R, and least on W. The middle-class pattern was similar with greatest relative improvement on Raven and Vernon of the British tests and least on Schonell and on the P.M.A. battery greatest relative gain on R, S, and N and least on V and W. This was not, however, statistically significant. The lower-class children show a significant shift in pattern with the greatest differences in the non-verbal mathematical area; the similar middle-class shift was not statistically significant.

IV.—Discussion and Conclusions.

The hypotheses derived from Ferguson's theory have been supported by the investigation. The patterns of abilities of lower-class children differ with educational experiences. In the grammar school verbal abilities develop most rapidly, in the technical school numerical, spatial and non-verbal abilities. The patterns of abilities of middle-class children were not affected by the type of school. The middle-class children in the technical school showed the same pattern of change as the lower-class children but in their case the differences were not significant. The number of middle-class children in the technical school was small and it may be that with a larger sample the differences would be significant.

These findings are relevant to the theoretical question raised earlier. Whatever the relative importance of constitutional and environmental factors, and of early cultural stimulation, later influences appear to play a significant role. Enviromental influences, specifically educational influences, affect the developing patterns of abilities if they provide a stimulus and make a demand that is new. Early influences may be lasting and generalized but, with this sample at least, they appear to be modifiable. The selective secondary school group are above average in ability and it may be that a stimulating environment is more effective with an able group. In other words, the relationship between such factors as constitutional factors, cultural stimulation outside school and educational experiences, and the development of ability may well be non linear.

These data do not indicate the precise nature of the changes that took place. The growth in verbal abilities of the lower-class grammar school group might simply be increased knowledge of grammar and facility in verbal expression. However, it is possible that the more significant changes of the kind described by Bernstein (1959) were taking place. The lower-class children may be acquiring new concepts and modes of thought as they acquire the formal language of the grammar school. Similarly, no attempt was made to distinguish between lower class children who did respond to the grammar school and those who did not. In general, the lower-class children did respond but the extent to which any particular child was affected by the school probably varied considerably. Jackson and Marsden (1962) plausibly suggested some of the home circumstances that might be relevant. Intellectual factors might be equally important.

Much recent work with culturally deprived children (Kirk 1958, Riessman 1962) has stressed the importance of early experiences and emphasized the importance of enriching the early environment. In practice this is difficult to do and many of the special facilities and programmes now being developed in the United States (Passow 1963; Burchill 1962; U.S. Office of Education, 1963) are for school age children and particularly adolescents. The Newsom report (Great Britain, Ministry of Education, 1963), which will presumably stimulate similar programmes in Britain, is concerned exclusively with secondary education. It may be that the earlier the work the greater the reward but these findings suggest that later influences can be expected to have some effect.

Perhaps the most important implications of this study are for the organization of secondary education. In the selective secondary schools lower-class children respond to the specific demands of the curriculum in a way that middle-class children do not. Bright lower-class children it seems, are offered an opportunity in the grammar school to make up for the lack of verbal stimulation in their earlier environment. Husen's study suggests that this is a function of the selective curriculum as contrasted with the comprehensive school, " Children from culturally less privileged homes responded more strongly to selective academic-type teaching " (Husen and Svennson 1960, p. 48). American studies indicate that the comprehensive school does not provide more equal access to higher education for lower-class children nor greater social interchange between children of different social classes than the traditional British selective secondary school. (Pohlman 1956; Stephens 1958; Coste 1959).

It is possible that the selective secondary schools are important for lower-class children but irrelevant to the development of middle-class children. The comprehensive school, as it exists, is not the ' open ' school described by Davis (1948) and Jackson and Marsden (1962). Comprehensive schools do not seem to be any more open to the interests and talents of lower-class children than the selective secondary schools. If able lower-class children are to continue to have the opportunities for higher education and for social mobility that has been offered by the grammar and technical schools, remedial, or more accurately compensatory, programmes like the New York Higher Horizons programme will probably have to be developed.

The differentiated curricula offered by selective secondary schools modify the patterns of abilities of lower-class children, but not those of middle-class children. The introduction of comprehensive schools would, therefore, be irrelevant for middle-class children but, unless some special provision is made, would impede the development of specialized skills by lower-class children.

V.—REFERENCES.

[1] BERNSTEIN, B. (1959). A public language. *Brit. J. Sociol.*, 10, 311-323.

[2] BURCHILL, G. W. (1962). *Work Study Programmes for Alienated Youth.* Chicago, Ill.: Science Research Associates.

[3] BURT, C. (1959). Class differences in general intelligence. *Brit. J. Stat. Psychol.*, 12, 15-37.

[4] CLARKE, A. M., and CLARKE, A. D. B. (1958). *Mental Deficiency.* Glencoe, Ill.: Free Press.

[5] CONANT, J. B. (1959). *The American High School Today.* New York: McGraw-Hill.

[6] COSTE, J. K. (1959). Some characteristics of high school pupils from three income groups. *J. Educ. Psychol.*, 50, 55-63.

[7] DAWE, H. C. (1942). Effects of a curriculum to develop verbal skills. *J. Experimental Educ.*, 11, 200-207.

[8] DAVIS, A. (1948). *Social Class Influences on Learning.* Cambridge, Mass.: Harvard University Press.

[9] DEUTSCH, M. (1960). *Minority Group and Class Status.* (Society for Applied Anthropol. Monog. No. 1). Ithaca, New York: Connell University Press.

[10] DOCKRELL, W. B. (1959). Socio-economic status intelligence and attainment. *Alta. J. Educ. Research*, 5, 16-22.

[11] FERGUSON, G. A. (1954). On learning and human ability. *Canad. J. Psychol.*, 8, 95-112.

[12] FERGUSON, G. A. (1956). On transfer and the abilities of Man. *Canad. J. Psychol.* 10, 121-131.

[13] FILELLA, J. F. (1960). *Educational and Sex Differences in the Organization of Abilities.* (Genet. Psychol. Monog. 61). Provincetown, Mass.: The Journal Press.

[14] FRASER, E. (1959). *Home Environment and the School.* London: University of London Press.

[15] GREAT BRITAIN, BOARD OF EDUCATION (1926). *The Education of the Adolescent.* Report of the Consultative Committee to the Board of Education. London: His Majesty's Stationery Office.

[16] GREAT BRITAIN, BOARD OF EDUCATION (1938). *Secondary Education with special reference to Grammar Schools and Technical High Schools.* Report of the Consultative Committee on Secondary Education. London: His Majesty's Stationery Office.

[17] GREAT BRITAIN, BOARD OF EDUCATION (1943). *Curriculum and Examinations in Secondary Schools.* Report of the Committee on Secondary School Examinations. London: His Majesty's Stationery Office.

[18] GREAT BRITAIN, MINISTRY OF EDUCATION (1959). *Fifteen to Eighteen.* Report of the Central Advisory Committee on Education. London: Her Majesty's Stationery Office.

[19] GREAT BRITAIN, MINISTRY OF EDUCATION (1963). *Half Our Future.* Report of the Central Advisory Council for Education. London: Her Majesty's Stationery Office.

[20] HEBB, D. O. (1949). *Organization of Behaviour.* New York: John Wiley and Sons.

[21] HUSEN, T., and SVENNSON, N. E. (1960). Pedagogic Milieu and the development of intellectual skills. *School Review*, 68, 36-51.

[22] JACKSON, B., and MARSDEN, D. (1962). *Education and the Working Class.* London: Routledge and Kegan-Paul.

[23] KENT, N., and DAVIS, D. R. (1957). Discipline in the home and intellectual development. *Brit. J. Med. Psychol.*, 30, 27-33.

[24] KIRK, S. (1958). *Early Education of the Mentally Retarded.* Urbana, Ill.: University of Illinois Press.

[25] MOSER, C. A., and HALL, J. R. (1954). The social grading of occupations, in, *Social Mobility in Britain.* London: Routledge and Kegan-Paul.

[26] OSBORN, R. C. (1947). How is intellectual performance related to social and economic background. *J. Educ. Psychol.*, 34, 215-228.

[27] PASSOW, A. H. (1963). *Education in Depressed Areas.* New York, N.Y.: Teachers' College, Bureau of Publications, Columbia University.

[28] POHLMAN, P. C. (1956). Relationship between abilities and socio-economic status and choice of secondary schools. *J. Educ. Sociol*, 29, 382-397.

[29] RIESSMAN, F. (1962). *The Culturally Deprived Child.* New York, N.Y. Harper and Row.

[30] STEPHENS, R. M. (1958). Stratification, education and occupational orientation, a parallel study. *Brit. J. Sociol.*, 8, 42-52

[31] U.S. OFFICE OF EDUCATION (1963). Bulletin No. 1, *Program for the Educationally Disadvantaged.* Washington D.C : Government Printing Office.

[32] U.S. CONGRESS (1963). *Education for all Children* Washington D C : Government Printing Office.

[33] VERNON, P. E. (1947). The variations of intelligence with occupation, age, and locality. *Brit. J. Psychol.* (STAT Section), 1, 52-63.

[34] VERNON, P. (1950). *The Structure of Human Abilities.* London: Methuen and Co.

[35] WARBURTON, F. W. (1958). Reply to "Measured intelligence is largely an acquired characteristic." *Brit. J. Educ. Psychol.*, 28, 291-293.

(*Manuscript received* 12*th June*, 1964).

THE VALIDITY AND RELIABILITY OF TESTS OF SPOKEN ENGLISH

By P. J. HITCHMAN
(*Department of Education, The University, Nottingham*)

Summary. Experimental tests in spoken English, comprising a ten-minute test for pupils in *prose reading* and *conversation* and a twenty-minute test for students in these two items and in *poetry-speaking* and *speech making*, took place in six training colleges and six secondary schools (173 pupils, 90 students and 15 raters taking part).

In one type of test each candidate was assessed on a single occasion by several raters, in another on two occasions, once by one rater and once by another, with ' equal and parallel ' tests. Performances were rated on various aspects, and on a *general impression*, of each item and on a *total impression* of the whole.

The results suggest · (1) That the types of speech test investigated have a reasonably high reliability and a reasonably high validity : i.e., that they are as valid and reliable as essay-type written examinations. (2) That these tests appear to be well-founded (i.e., that they comprise elements on which satisfactory judgments can be made). (3) That fatigue does not affect the reliability of raters' judgments.

I.—Introduction.

Tests in spoken English in the past have been concerned mainly with *prose reading* and *poetry speaking*. Radio and television have stressed another, and perhaps more important, side of human communication : that of people addressing audiences or being questioned or talking together. As a result, modern tests tend to provide a balance of two sorts of speaking, on the one hand Oral Interpretation (as in *prose reading* and *poetry speaking*) and on the other Oral Composition (as in *speech making* and *conversation*). Examinations involving subjective judgments raise the very large questions of validity and reliability. Little is known about the validity and reliability of tests in spoken English.

II.—Outline of previous research.

The reported research is meagre ; almost all of it is American and most of it is in the field of public speaking. A large part has been done with student judges. Monroe, Remmers and Lyle estimated, for public speaking, that the reliability of the judgment of a single student was 0·3 and that of a single instructor 0·5 (Thompson, 1943). Other investigators, comparing the judgments of students and trained instructors in this field found correlations ranging from 0·34 to 0·85. Leyden found it necessary for the rating of recordings of speeches to have as many as thirty-six student judges in order to secure a reliability coefficient of 0·90 or better (Auston, 1953).

In addition to the work on public speaking several American researchers have enquired into the reliability of rating scales requiring judgments to be made covering particular aspects of speech-making (e.g., attitude, language, fluency, organisation, etc, in number from seven to sixteen aspects). Bryan and Wilke (1941), Eckert and Keys (1940), Greene (1941), Knower (1929) and Gilkinson and Knower (1941) produced a wide range of reliability coefficients measuring the agreement of two or more raters. Perhaps the most important research into public speaking is that of Becker (1962) who, using a scale of eleven elements, found that raters could discriminate only three : an analysis-content factor, a delivery factor and a language factor, the last appearing to be the most independent.

In the field of oral interpretation Weaver (1924) tested the 'between group' reliability of the composite judgment of small groups of students and found a correlation of 0·9. Knower (1940) reported that the pooled assessments of sixteen trained judges would be required to produce a similar coefficient. Seedorf (1949) found that any group of judges numbering forty can be expected to agree fairly closely in ranking readers.

In American research into the reliability of various forms of speaking, including speech-making, two of the elements of presentation that have been freely used are *voice* and *articulation*. Average intercorrelations for single judgments of these elements are only moderate but group judgments show considerable reliability.

It is the opinion of American experts that the importance of the actual rating or ranking scale used is quite secondary to the quality of the assessor employed (Auer 1959, Thompson 1943 in the field of public speaking, Seedorf 1949 in that of oral interpretation).

There appear to be no records of test validation in the field of spoken English either in Britain or America.

The scanty evidence available suggests that the reliability of the judgment of trained judges when judging singly in the field of public speaking and when judging singly *voice* and *articulation* is rather low. Since the judges and subjects in the experiments reported were known to each other it is probable that in experiments in which they were not known to each other the reliability would be even lower.

III.—PLAN OF THE INVESTIGATION.

(i) *Subjects and Materials.*

Two types of test were devised : Type A, in which students and pupils were examined singly by the writer but were *assessed* by three raters (in two tests) or four (in one), sitting as observers in the background and not in touch with each other ; and Type B, in which each student and each pupil was examined twice with 'equal and parallel' tests, once by the writer and once by one of a number of other assessors.* Seven raters, thirty students and twenty-seven pupils took part in the A tests held in two colleges and one school, and eight raters, sixty students and 146 pupils in the B tests held in four colleges and five schools.

Each test covered a working day of five hours' examining. For each subject the college test lasted twenty minutes during which he or she was examined in prose reading, poetry speaking, speech making and conversation (fifteen students). For each subject the school test lasted ten minutes, the test elements being prose reading and conversation (approximately thirty pupils). The aspects or ' details ' of each item to be rated were :

---

* Medley and Mitzel distinguish between the two forms of 'reliability' measured by the Type A and the Type B tests. The correlation between scores based on observations made by different observers at the same time (as in the Type A tests) they call a *coefficient of observer agreement*. The correlation to be expected between scores based on observations made by different observers at different times (as in the Type B tests) they refer to as a *reliability coefficient*. The former, they say, tells us something about the objectivity of an observational technique, but only the reliability coefficient tells us how accurate our measurements are (Gage, 1963).

Prose reading : Voice, diction* (articulation), interpretation, general impression.
Poetry speaking : Ditto.
Speech making : Voice, diction, power of verbal expression, manner and delivery, plan and quality of content, general impression.
Conversation : Voice, diction, power of verbal expression, general impression.

(ii) *Experimental Procedure.*

American speech experts have noted the importance—and the difficulty—of attaching exact and intended meanings to the terms on the rating scale (Wilke 1935, Knower 1944, Auer 1959). With the help of three other assessors who took part in the experiment, a rating scale was designed to be as objective as possible in its definitions of the aspects of speaking to be observed. The writer tried to ensure that the assessors were as nearly of one mind as possible by careful and rigorous briefing, including discussion of questions they raised.

A problem concurrent with that of description was the choice of a method of scoring. Thompson's work suggested that no one rating technique is vastly superior to any other (1943). Symonds (1925) had found ***ranking*** and ***rating*** equally reliable, and recorded that " everyone I have asked tells me rating is more pleasant than ranking " (p. 190). The writer chose rating. It was decided to use a numerical eleven-point scale, 0–10, and to define only the ends of the continuum, 9 and 10, 0 and 1. The ' detail,' ' power of verbal expression,' was set out thus :

10, 9 :	Spontaneous and fluent command of language. Well-ordered arrangement of ideas. Evidence of ability to develop a theme. Subject matter of good quality. Vocabulary suitable and of adequate range. Idiomatic use of English.
8, 7 :	
6, 5, 4 :	
3, 2 :	
1, 0 :	Finds it difficult to say anything : or is verbose. Ideas are muddled. Finds it difficult to develop a theme. Fails to keep to the point. Inadequate vocabulary. Uses slang inappropriately.

A full-scale preliminary experiment to discover any practical difficulties in the testing procedure was carried out in the first of the colleges. Every effort was made in this and the succeeding tests to standardise conditions. Items were always in the same order : prose, poetry, speech, conversation. Each candidate was asked to stand for reading at a point (marked on the floor) eighteen feet from the centre of the assessor's chair. For the conversation element of the test a chair was placed for the candidate at such a distance and such an angle from the desk that the centre of the candidate's chair was five feet away diagonally from the centre of the examiner's chair, with one end of the desk intervening.

---

* The term *diction* was preferred to *articulation* since the writer felt that to most assessors in this country it would suggest the *total effect in running discourse* of the making of consonants and vowel glides, whereas the term *articulation* suggests rather the narrower application to short segments of speech with a closer examination of the relationship of individual plosives, fricatives, nasals, etc., to preceding and following vowel sounds.

There were four parts to the test for students, who would need a brief rest between items. After the *prose reading* the candidate would be asked to sit (at the eighteen feet mark) and glance at his poem for a moment : and after the *poetry-speaking* he would be asked to sit (at the same distance) and look over the heads of his prepared talk.

IV.—Results.

Validity.

The only external criterion that can be used for the assessment of the validity of tests in spoken English is the estimate (based on long knowledge of the candidate) of the lecturer or teacher. Over the whole range of tests lecturers and teachers expressed their valuations variously, from a four-grade scale A—D to a composite assessment ranging from (a theoretical) zero to a maximum possible of 200. It was decided that the best statistical technique for comparing criterion assessments with those of raters was to set up a series of 2×2 tables, each resulting in a tetrachoric correlation r_t arrived at by using Thurstone's Computing Diagrams (Cheshire, Saffir, Thurstone, 1933).

The samples used in the tests were small, thirty-one or fewer (fifteen only in college tests). This means that, if r_t coefficients of validity are to be significant *standing by themselves*, they must be very high. However, Yule and Kendal (1953, p. 473) have suggested that where there is a number of tables of similar data from different fields, it is permissible to obtain a value of *P* for the aggregate (*chi squared* being additive). The 2×2 tables were tested for significance by the *chi squared* method (with the Yates' Correction for Continuity) and Fisher's Exact Test (1946).

Two groups of validity coefficients for individual raters and ' raters combined ' in the A Tests and the B Tests separately were found. The college or school assessment was compared with (1) the ' Total Impression ' marks and (2) the ' Sum of Details ' marks. Mean coefficients ranged from 0·66 to 0·83 ($P < \cdot 01$). These are considered satisfactory.

If the top 75 per cent. of candidates are regarded as passing a Spoken English test and the lower 25 per cent. as failing, a contingency table can be prepared (Daniels, 1951 and 1954) showing the relationship between the estimates of performance of a candidate-list of 100 made by lecturers in colleges and teachers in schools and those made by external assessors.

TABLE 1

Contingency Table for r_t 0·8, Showing the Relationship Between the Assessments of Colleges and Schools and those of ' Assessors Combined ' (Type A Tests).

		Estimates of Colleges and Schools		Total
		Top 75%	Lower 25%	
Estimates of ' Assessors Combined '	Top 75%	67	8	75
	Lower 25%	8	17	25
	Totals	75	25	100

This table shows that there is a disagreement on placing in the pass-list of $8+8=16$. On the average eight candidates who would be forecast as passing the test by the college or school would be failed by the examiner and eight who would be forecast to fail would be passed by the examiner; a misclassification of 16 per cent.

The ' Assessors Combined ' r_t for the Type B tests is for the ' Total Impression ' and the ' Sum of Details ' approximately 0·7. In this case misclassification would probably be, on the average, 20 per cent.

Reliability.

The general consistency of markers is shown in Table 2.

TABLE 2

MEAN INTERCORRELATION OF VARIOUS MARKS OF FOUR, THREE AND PAIRS OF RATERS IN THE A AND B TESTS (P<·01 in all cases).

Type of Test	N	Total Impression	Sum of Two General Impressions	Sum of Four General Impressions	Sum of Six Detail Marks excluding General Impressions	Sum of Fourteen Detail Marks excluding General Impressions
A						
(1) One College Four Raters	15	0·71		0·83		0·83
(2) One College Three Raters	15	0·74		0·78		0·77
(3) One School Three Raters	27	0·77	0·72		0·75	
B						
(1) Four Colleges Pairs of Raters	60	0·72		0·75		0·79
(2) Five Schools Pairs of Raters	146	0·69	0·76		0·72	
Total N	263					

These figures represent a moderate agreement between assessors. There is a rather surprising ' between-test ' agreement.

A comparison of the mean intercorrelation of ' General Impression ' marks and of ' Sum of Details ' is of interest.

As schools usually take only prose reading and conversation (of the four items treated in this article) Table 4 is presented to show intercorrelations representing the consistency of agreement of assessors when the marks for these two items are combined.

The figures in Tables 2, 3 and 4 represent roughly the same degree of inconsistency in marking. (This is probably no more than is to be found in the marking of written English composition in schools). If we take a rough average of the figures in these tables and set up a 2×2 Contingency Table (as in Table 1) we can say that misclassification of candidates (on the same basis of Pass/Fail) would probably be, on the average, 18 per cent.). It will be seen from Tables 2, 3 and 4 that the reliability of the ' General Impression ' mark for each item is very little lower than that for the ' Sum of Details.' This is almost certainly

TABLE 3

Mean Intercorrelation of ' General Impression ' Marks and of ' Sums of Details ' for Four, Three and Pairs of Raters, in Both A and B Tests (P< ·01 in all cases)

Type of Test	N	Prose Reading		Poetry Speaking		Speech Making		Conversation	
		General Impression	Sum of Three Details	General Impression	Sum of Three Details	General Impression	Sum of Five Details	General Impression	Sum of Three Details
A									
(1) One College Four Raters	15	0·80	0·81	0·82	0·85	0·70	0·81	0·76	0·80
(2) One College Three Raters ..	15	0·67	0·70	0·72	0·67	0·74	0·78	0·67	0·68
(3) One School Three Raters .	27	0·65	0·68	—	—	—	—	0·72	0·77
B									
(1) Four Colleges Pairs of Raters .	60	0·79	0·80	0·70	0·79	0·61	0·72	0·60	0·68
(2) Five Schools Pairs of Raters	146	0·69	0·70	—	—	—	—	0·63	0·63
Total N .	263								

TABLE 4

Mean Intercorrelation of Various Marks of Pairs of Assessors of ' Details ' (for Prose Reading and Conversation only) for All Schools and Colleges in Both A and B Tests (P< ·01 in all cases)

Type of Test	Two General Impressions Prose and Conversation	Six Details Prose : Voice, Diction, Interpretation. Conversation Voice, Diction, Power of Verbal Expression	Eight Details (as for Six) Plus Two General Impressions
A	0·757	0·782	0·786
B	0·753	0·744	0·766

due to the fact that it has been steadily built up in the mind of the assessor as he has been rating the particular ' details.' The practical value of this is that, if marking is done in this way, a set of ' General Impressions ' is likely to prove just as reliable a set of scores as a ' Sum of Details,' and it is much less cumbersome for assessor and examining authority.

Assessor Fatigue.

It is possible that testing for a whole day may, in spite of the breaks arranged for coffee and lunch, give rise to some fatigue effect in the assessor. Examination of the scores for fatigure trends was done in two ways:

(1) Comparison of ' Sums of Scores ' for consecutive candidates.

In this method the hypothesis is set up that, when an assessor is tired, he will tend to award higher marks than when he is not tired. If this is so, then high marks will tend to appear at some time during the afternoon, probably towards the end.

The scores were divided into four blocks covering the working day, two before lunch and two after. In two of the A tests there was, indeed, a block of high scores at the end of the day in the case of six of the seven assessors involved. But in the third of the A tests, involving three of the six assessors mentioned, there was a block of high scores in Period III. Since the pattern found in the figures for this third test is a different one from those for the other two tests, and since in any of the Type A tests all the assessors during any one period of time would be assessing the same sequence of candidates (which would not be the case in the Type B tests), it would appear that the variations arise from genuine differences in the students in these blocks. The hypothesis is also unsupported by the results of analysis of the Type B tests, in which the number of highest ' Total Scores ' for each of the four periods of the day came to: I 4, II 7½, III 1½, IV 5.

(2) Comparison of ' Sums of Differences ' between scores for consecutive candidates.

A second form of analysis is to take each ' detail' in turn, sum the differences between the scores for consecutive candidates, and compare the sums of all ' details ' for the four periods of the day (starting to count afresh after lunch). The hypothesis is that a period during which the assessor's attention may wander (e.g., that after lunch, or Period IV when he may be tired at the end of the working day) may be expected to produce scores for each ' detail ' with smaller variations between consecutive candidates (his score-sheet showing at a glance the marks he awarded to the preceding candidates) than a period during which he may be expected to be alert (Periods I and II).

The four assessors working in the first of the Type A tests produced a period of lowest ' Sums of differences ' immediately before lunch, Period II. The three assessors working in the other A tests produced a similar result in Period III. Again, it would seem that the variations arise from genuine differences in the students in these blocks, more especially as the results are unsupported by those of the Type B tests, in which the number of lowest ' Sums of differences ' came to: I 9, II 2, III 2, IV 5.

Considering the results of the A and B tests together, one may say that it has not been possible by either method of examination of scores to discern any fatigue effects in assessors.

V.—DISCUSSION.

Reliability intercorrelations between assessors fell roughly between +0·7 and +0·8. The main sources of unreliability are likely to be the assessor, the candidate, the system of rating, the test-material and the conditions in which the test is held. In this experiment one source of unreliability was certainly the assessor. The fact that for the Type A tests the correlations were below +1·00 could not have been due to variation in the performance of the candidate

261212

or in the test-material (there was a single performance and only one set of material) ; it must have been due to disagreement in estimation by the assessors. The mean correlations for the Type A and the Type B tests were roughly the same. Since, in the case of the Type B tests each candidate was examined on two occasions, once by each of two assessors, in different rooms with different test-material, any of the sources mentioned could have been responsible for the element of unreliability. It is reasonable to infer that the ' unreliability element ' present in the assessors in the Type A tests was also present in those in the Type B tests ; the nature of the test was such that an ' unreliability element ' could also have been present in candidates, rating system, test-material and test conditions.

An immediate improvement in reliability could be obtained by employing two or more assessors in any one test and pooling their results. Much can be done in the standardisation of test conditions. Research is necessary to decide how, if at all, unreliability can be eliminated or diminished in the test material.

In spite of the fact that correlations fell considerably below +1·00, we may yet contrast the reliability of the marking of the single assessor in the spoken English test not unfavourably with that of examiners in other subjects to which it may reasonably be compared (i.e., subjects which require essay-type answers). Daniels (1954), for example, found an average correlation (*between* markers) of +0·779 for the marking of essay-type answers of typical G.C.E. ' O Level ' Geography papers. When we consider that the tests are also as valid as tests requiring essay-type written answers* we can say that with only one examiner the test in spoken English is justified.

The experiment seems to show that (within the limits of the study) the testing of spoken English can be done satisfactorily (statistically and psychologically speaking) provided certain conditions are observed. These are :

(1) The assessors must be competent and carefully briefed.

(2) They should be instructed to assess first a set of ' details ' (such as the sets used in the present study) and finally for each item a ' General Impression ' (the latter to be assessed quite separately from the other ' details ').

(3) Assessors should return to the examining authority a set of ' General Impressions,' one for each item in the spoken English test.

(4) A mark-sheet should be used which allows the assessor, when he is judging a candidate, to see at a glance just what grades he has allotted for any one ' detail ' to all preceding candidates.

(5) The nature and purpose of the tests and the testing procedure should be explained carefully to candidates before the tests take place.

---

* Valentine (1959, p. 649) states that between the orders of a sixth form of, say, thirty boys in English and History in the old School Certificate examinations and the order of merit drawn up by the teacher in the same class r=about 0·65 on the average of many schools.

VI.—REFERENCES.

AUER, J. J. (1959). *Introduction to Research in Speech* New York Harper and Bros.

AUSTON, J. (1953). Methods and levels of measurement in speech *Educ. and Psychol. Meas.*, 13, 228-247.

BECKER, S. L. (1962) The rating of speeches · Scale independence. *Speech Monog*, 29, 38-44.

BRYAN, A. I., and WILKE, W. H. (1941). A technique for rating public speeches. *J of Consult. Psychol.*, 5, 80-90.

CHESHIRE, L., SAFFIR, M., AND THURSTONE, L. L (1933). *Computing Diagrams for the Tetrachoric Coefficient.* Chicago · University of Chicago Bookstore.

DANIELS, J. C. (1951). The interpretation of correlations. *Bulletin of the University of Nottingham Institute of Education*, 7, 1951.

DANIELS, J. C. (1954). Testing geography at the ordinary level of the general certificate of education. *Brit J. Educ. Psychol.*, 24, 180-189.

ECKERT, R. G., and KEYS, N. (1940). Public speaking as a clue to personality adjustment. *J. of App. Psychol.*, 24, 144-153.

FISHER, R. A. (1946). *Statistical Methods for Research Workers.* London · Oliver and Boyd.

GAGE, N. L. (1963), ed. *Handbook of Research on Teaching.* Chicago : Rand McNally.

GILKINSON, H , and KNOWER, F. H. (1941). A study of standardised personality tests and skill in speech. *J. of Educ. Psychol*, 32, 161-175.

GREENE, G. S. (1941). The correlation between skill in performance and knowledge of principles in a course in speech-making *J. of App. Psychol.*, 25, 232-242.

KNOWER, F. H. (1929). A suggestive study of public speaking rating scale values. *Quart. J. of Speech*, 15, 30-41.

KNOWER, F H. (1940). A study of rank order methods of evaluating performance in speech contests. *J. of App. Psychol.*, 24, 633-644.

KNOWER, F. H (1944). What is a speech test? *Quart. J. of Speech*, 30, 485-493.

SEEDORF, E H. (1949). An experimental study in the amount of agreement among judges in evaluating oral interpretation. *J. of Educ. Res.*, 43, 10-21.

SYMONDS, P. M. (1925). Notes on rating. *J. of App. Psychol.*, 7, 188-195

THOMPSON, W. (1943). *An Experimental Study of the Accuracy of Typical Speech Rating Techniques* Unpublished thesis Evanston, Ill. : North-Western University.

VALENTINE, C. W. (1959). *Psychology and its Bearing on Education.* London · Methuen.

WEAVER, A. T (1924). The prediction of talent for vocal expression in reading. *J. of App. Psychol.*, 8, 159-186.

WILKE, W. H. (1935). Subjective measurements in speech : a note on method *Quart. J. of Speech*, 21, 53-59.

YULE, C., and KENDAL, M. (1953). *An Introduction to the Theory of Statistics.* London : Griffin.

(*Manuscript received 4th January*, 1965.)

TOWARDS A RATIONALE FOR TESTING ATTAINMENTS AND ABILITIES IN AFRICA

By S. H. Irvine*

(*Formerly Lecturer in Education, University College of Rhodesia and Nyasaland*)

Summary. A brief review of recent research in Africa in the field of cognitive measurement is followed by a discussion of the central issue of test score meaning in testing educational skills and aptitudes. Influences on test score meaning are hypothesised, among them sex differences and role relationships. A major influence is the quality of school attended Two factor analyses are reported in an attempt to clarify the application of traditional constructs in the mapping and prediction of educational skills. They indicate that constructs like ' g,' ' v ed.' and ' n ' may have legitimate psychological meaning in non-western cultures. To assess the comparative validity of the same test used on different cultural groups, item difficulties for Raven's Progressive Matrices between British and African groups were correlated. The resulting rank correlation (·695) was taken as *prima facie* evidence that the test scores were only partly comparable, but encouraged using the Raven as a ' g ' reference axis in factor analysis.

The results quoted were gained from a representative sample of over 1,600 African eight-year educated children in Southern Rhodesia, and from a follow-up group of 291 from the same sample in Form I (nine-year educated). In addition, concurrent validities with primary school headmasters' estimates, using combinations of traditional and experimental tests, were in the region of ·90 while multiple correlations of ·70, uncorrected for range restriction, were gained with a Form I multiple criterion. These correlations were achieved by selecting variates for regression analysis with reference to the factor analyses carried out, in an attempt to check the psychological meaningfulness of the factor solutions.

I.—Introduction and review of research.

With the emergence of independent states in Africa has come the realisation that local people would have to take over jobs previously held by expatriates. Middle and high level manpower has had to be identified, educated and trained, often in a hurry, to meet the deadline of independence and to avoid an administrative and technical vacuum in the years immediately following self government. In latter years selection mechanisms have been employed in industry and commerce and in the schools at the end of the primary stage with the main purpose of assessing the attainments and abilities of Africans in behaviour areas that are western in origin. People are required who have the verbal and technical skills appropriate to the running of a modern state. Furthermore, such people compete with their peers and are seldom, if ever, assessed along with or compared with others of different ethnic and social backgrounds, except where there are relatively large numbers of Europeans living in industrial areas of independent African states such as the Copperbelt in Zambia.

Recent studies in Africa have, therefore, been truly predictive in nature, whereas early work was essentially comparative in its aims. The researches of Loram (1917), Oliver (1932, 1933, 1934), Nissen and Machover (1935), Fick (1939) and even Biesheuvel's *African Intelligence* (1943) showed more awareness of the differences in the approach to cognitive problems between races and attempts to explain them than the need to predict African success or failure in a given task. The watershed between comparison and prediction was reached in a little publicised but highly significant study made by an Army team of psychologists working under Col. A. Macdonald (1944-45) in the Middle East and in East Africa. In a year Macdonald had adapted and constructed no fewer than fourteen measures of the aptitudes required for skilled tasks in the modern army. The results, despite criterion instability, were promising both for performance and paper and pencil tests used among semi-literate African recruits. Macdonald and his team discovered, for the first time, that the presentation of such tests to Africans was an important part of the key to their validity. Special methods of demonstration and explanation were devised in English and Swahili, the East African lingua franca. Macdonald also subjected his tests to factor analysis, as reported in Vernon (1961a) and began the study of psychological constructs in Africa ; but this will be taken up later in the paper. Biesheuvel (1952, 1954) reports the work of the team at the National Institute for Personnel Research (N.I.P.R.) Johannesburg. This team constructed the now famous *General Adaptability Battery* of performance tests for the screening of illiterate African mine workers ; and again emphasis was put on the demonstration of the correct techniques of taking the various sub-tests by film and mime, with commentary in a polyglot vernacular. Paper and pencil tests for fully literate Africans began to be used in predictive studies in Zambia in 1956, with Clark's investigation of secondary school selection procedures. Brimble (1963) constructed an abstract reasoning test soon afterwards for the same purpose. By 1960, governments in various African territories showed interest in classifying pupils according to their academic skills by sending delegates to and supporting the meetings of CCTA/CSA (Council for Technical Cooperation in Africa South of the Sahara) at which psychologists from the continent of Africa met and discussed ways in which tests could be used to improve selection procedures ; and an important conference held in Accra in 1961 resulted in the publication under Taylor's editorship in 1962 of *Educational and Occupational Selection in West Africa.* From these meetings began the most recent phase of test development in Africa. Teams under the leadership of Schwarz (1961) of the American Institute for Research based in Lagos, and Roberts of the N.I.P.R. in Johannesburg began independently to construct group tests for use in the upper primary (eight years of education) stages. In Rhodesia the University College undertook research at the end of the primary school, using the N.I.P.R. tests as the best then available and combining them with traditional tests to discover the most efficient combination of both for secondary school selection. By 1963 the A.I.R. tests constructed by Schwarz in Lagos had become available for trial outside Nigeria. Consequently both the Schwarz tests and the N.I.P.R. tests for which Roberts was responsible were tried out in the *Northern Rhodesia Mental Ability Survey* (MacArthur, Irvine and Brimble, 1964) which ran in parallel with the study in Southern Rhodesia. Silvey (1963) in Uganda was also investigating the use of psychological tests in the eighth year of education for senior secondary school selection, so that the period 1960 to 1964 may be marked as one of intensive study of the problems of administering tests to African schoolchildren assumed to be literate in English. Also, the considerable contributions of the Belgian psychometrists in the Congo prior to the independence of that country deserve more recognition. Work in French-speaking Africa and in Central and East Africa has benefited from the various researches carried on at the Centre D'Orientation Professionelle in Leopoldville, particularly that involving experiments with Raven's Progressive Matrices (1958). One must not forget, either, the studies of Want-man (1962) in Malaya and Vernon (1961b) in Jamaica as evidence of a wider interest in the application of psychological tests to non-European populations.

---

* Now Lecturer in Education, The Institute of Education, The University of Bristol.

II.—INFLUENCES ON TEST ADMINISTRATION AND SCORE MEANING

In nearly all of these recent investigations, the basic problem has been to discover how to use psychological tests with groups of pupils who have never seen their like before. Macdonald and Biesheuvel tackled the task by demonstrating and using a common language. Our own procedures in Southern and Northern Rhodesia did not differ very much in kind from the early studies. They differed in degree in that the learning task involved in answering questions of a multiple-choice type is, for unsophisticated subjects, a complicated one. Again, the sort of question set in such tests was often new to the pupils. Each different type of question had to be demonstrated and examples given. Early experiments with *Raven's Progressive Matrices* in the Rhodesias showed that variance could be introduced or excluded at will by altering the form of test presentation and that validity changed with different presentation. Silvey, in Uganda, found the same thing with similar populations. We concluded that the format of the test carried with it its own source of variance, as Vernon (1958) showed in his monograph on tests of English comprehension, and that the immediate transfer of what had been learned about different types of items in practice to the actual test situation carried with it another important individual or group variance. Individual ability to transfer the cues and codes of items encountered in practice varies, but groups themselves vary with the quality of the test administrator. Hence, there may be elements in a test situation that confound attempts to gain a relatively pure measure of the ability being tested.

If, however, one were able to control the test situation by simply ensuring that questions and examinations were only of the type that the schools and the teachers had prepared the children to answer, one would still find in Africa chance group influences overlaying individual differences. For example, it was possible in Rhodesia to obtain normalised estimates of school success in examinations both in the primary school and in the secondary school. These estimates were based on past examination success and, independently, on the degree of pre-selection undertaken by the primary schools. These two estimates and the result of a single intelligence-type test, when correlated with the criterion of performance at the end of Form I in the secondary schools gave a multiple correlation of about ·5 (N=291). School quality was, in Southern Rhodesia, one of the most powerful influences and it blanketed individual differences considerably. Educationally, a wide range in the standard of teaching offered at schools is a bad thing. For predictive purposes it offers a cheap and useful, if non-static, element in the selection procedure especially if measures can be found that are relatively uncorrelated with school quality.

The test constructor in Africa must, therefore, take particular note of the following main influences on individual test scores in educational measurement:

(1) The form in which the test is administered.
(2) The types and variety of items in the test.
(3) Local school standards.

This aspect of research in Africa has in a sense brought into sharp focus some of the simple rules of test construction because most large-scale testing in the highly developed countries began after universal primary education was the norm and not the exception. Mass media, rising standards of living and the supplements to reading readiness in the form of picture books, toys and children's newspapers have brought about, in conjunction with high teaching standards, a relatively (by world criteria) stable and homogeneous social and educational climate for learning in western societies. Individual differences are the more easily traced and defined since group differences are small. In Africa, the school is often the only point of culture-contact and for virtually all of the children whose parents are subsistence farmers (some 90 per cent. of the population) environmental supports to learning are not possible. Individual differences are obscured : and as a consequence test users in Africa are acutely aware of problems that may seldom have confronted testers in western societies because conditions did not exist to highlight them to quite the same degree as they do in Africa.

Turning now to a more technical and practical matter, as a final example, one encounters in Africa the problem of age in the technique of test standardisation and norming. In Britain, all births are registered by law and chronological age is a verifiable constant within very fine limits. In Africa, only rough estimates of age exist, and these are usually tailored to the requirements of educational system. In Southern Rhodesia there may be, at the end of the primary stage, a large age range within any one standard or grade of education. When reputed age is correlated with attainment or ability tests there is a small negative and significant correlation between age and test score. Norming by such ages and using 'accepted' standardisation procedures would serve to penalise the younger children, so that some other method must be found. At present it is suggested, because of the difficulties of assessing age, that the tests will have to be normed within standards or years of completed education and allowances given to children who have accelerated through the system, thus penalising repeaters.

Although age is not verifiable we may be thankful that sex is an attribute that very seldom undergoes any significant change. Testing at the age of transfer in Britain has indicated that girls do slightly better in verbal tests than boys. Exactly the opposite holds in Central and East Africa, on the whole. Males have a verbal advantage in all tests where English is the medium, but in abstract tests there are very few significant differences between boys and girls. This result may be attributed to the social mobility of males in Central African societies, to the increasing social dominance achieved by such mobility, even in matrilineal societies, to the deference of females to them in conversation, and possibly to the encouragement to succeed given by a predominantly male teaching staff to boys rather than girls. Males are expected to learn English and learn it well, since success in school and in skilled employment depends on mastery of it. Males did not have it all their own way, however, during our experimental pilot-study work with *Morrisby's Compound Series Test* in Salisbury with a random school sample of just over 200 pupils. Morrisby (1955) claims that in Britain this test has a high 'g' loading and is relatively unaffected by school attainments and general environment. Users of this group test will remember that it is essentially one of bead stringing where the beads differ in shape, size, colour and position. Not surprisingly, because of their experience in the locally taught craft of bead-work, girls did significantly better than boys in our sample. There are two conclusions from this result. First, environmental supports for particular cognitive exercises produce group differences and, more important, one could not say that the test, for this African sample, was "a measure of pure intelligence," as Morrisby (op. cit. p. 1.) claims it is for British children. The meanings of two test scores, one British, the other African psychologically are different because of different environments. If one were using the same test to predict success on a common criterion of, say, work performance, then as far as evaluation among either European or African peoples were concerned, the test score could be held to have finite and common meaning. This *predictive* meaning should be held separate from *descriptive* or environmental meaning (Irvine 1963, p.45). Predictive meaning allows value judgments when the criterion is itself recognised as a value judgment. Test scores taken by themselves and group averages based on such scores (compare Shuey (1958), whose work is based on such techniques) are not sufficient proof of the superiority or inferiority of one group or another unless one is certain that the sources of variance, and, more important, the proportions of each definable source of variance, within the test situation are the same for both groups and that, as a consequence, the test scores for the two groups have the same psychological meaning.

III.—COMPARATIVE AND CONSTRUCT VALIDITY

So far, this paper has suggested that test content and format, school conditions, age, sex and its accompanying role expectation affect scores in Africa to an extent and in a manner that would prevent detailed psychological comparison between European and African groups. If this is so, then one must ask if the frame within which human abilities have so far been described is applicable to results from African populations. Silvey (1963) has rightly said that at this stage in the history of testing the predictive or evaluative meaning of test scores is primary and that construct or content validities are secondary in Africa. He further stresses that factor analytic techniques describe tests better than they describe the mental processes which produce the test answers. There may be ways, however, of checking on whether a test has *comparative* validity from group to group and one of these may partly consist in the correlating of item difficulties derived from the same test with different sub-populations. Studies with Raven's *Progressive Matrices* by Vernon (1950), Slater (1948) and Keir (1949) showed that item difficulties for the Raven test correlated to approximately ·95 among sub-groups, children and adults in Britain. It is thus arguable that in a test of this nature the test was difficult for these groups for the same reasons. The Raven has a high ' g ' content and we might, therefore, further argue that tests sharing a reasonable degree of item-difficulty agreement are testing the same things and that the groups to which the tests are applied are reacting in the way that the test constructor hoped they would. The correlation between item difficulties in Keir's study of British school-children and those calculated for the same test on a representative sample of over 1,600 African eight-years educated pupils in Rhodesia was ·695. With only 50 per cent. of variance between item sets accounted for, there is ample room for an interpretation of the Raven test score from the African sample as being not exactly equatable in meaning with that of the British group ; and to argue that this is so because the items are not perceived in the same way by both groups ; but enough of a relationship to suggest that the construct being tested may be shared by, although may not be identical for, each group.

This line of thinking was used (Irvine, 1964) as a criterion for rotating axes in two factor analyses designed to select variates from a large number to be used for psychologically meaningful prediction of secondary school performance for the sample just mentioned. The Raven was hence taken as a ' g ' reference test and an axis passed through this and another abstract test that occupied a position close to it in factor space. Consequently, the Principal Factor solutions, first rotated by the Varimax method, were adjusted by using this reference technique. The results were, for the primary school sample of 1615 non-repeating children, an eight factor solution involving twenty-six variates for which 48 per cent. of the whole variance was accounted for. The main factors appeared to be ' g ' (23 per cent. of common variance), second language v.ed. (23 per cent. of common variance) goodness of school (17 per cent. of common variance), ' n ' a number factor (10 per cent. of common variance) and a general knowledge or social awareness factor (7 per cent. of common variance). These factors appeared to be psychologically meaningful, the others were an artefact of family size and position and two descriptive factors that were the product of certain environmental variables. The second analysis, performed on the 291 follow-up cases and with twenty-eight variates in Form I, all that remained of the original sample, showed similar factors of ' g,' school quality, second language v.ed., general knowledge, number and what was termed male educational aptitude.[1]

As a test of the factor hypotheses, variates were selected, using the factor loadings as guides, in an attempt to produce economical and psychologically meaningful regression on concurrent and on follow-up criteria, which were themselves included in the factor analyses. Good fits for various regression equations were found by using the variates selected from the final factor solutions and these were checked by allowing free step-wise regression among all possible variates. Correlations of ·90 with concurrent primary headmasters' estimates were possible and with Form I performance a multiple correlation of ·7 was achieved. This, allowing for restriction of range, could be interpreted for the whole population as a correlation of ·89, using a combination of traditional tests, experimental tests, headmasters' estimates and assessments of school quality.

We concluded that western frames of reference could be used to analyse data derived from African populations and that interpretation of factor solutions in terms of western constructs did not invalidate regression. Our conclusions were borne out by previous factorial studies undertaken in Africa. Macdonald performed a factor analysis on his African recruit data and this is reported, unrotated, by Vernon (1961a p.106). Rotation of this analysis showed that a matrix test, adapted from Raven's *Progressive Matrices* had the highest loading on the first factor which was postulated as ' g ' and that two other factors appeared to be mechanical aptitude and manual dexterity. Biesheuvel's (1954) analysis of the General Adaptability battery also produced a large general factor of ' general adaptability.' McFie's (1961) paper on testing and re-testing African trainees showed a general factor on first testing that was reduced on re-testing two years later after training with more emphasis on group factors. Brimble's (1963) factor study of his test again showed a large general factor and a smaller group space-perception factor. These previous studies have tended to demonstrate that general factors are the norm among culturally heterogeneous African populations. Our own work also indicates that a conglomerate general factor that may not be 'g' could arise among populations where many group influences overlay individual differences (Irvine 1964, pp. 219 *et seq.*). By using a well-tried test as fixed point for a reference vector this conglomerate was split into second language v.ed. and ' g,' which is consistent with our hypothesis that gross differences in educational opportunity will produce a flux between which is traditional, learned and not wholly dependent on ability and what is relatively new, hence more dependent on individual ability in African schools. What is called intelligence may, then, in Africa, be at this point in social history, far more a product of education than it is in western societies and the variance in an intelligence test score may owe more to environmental influences, at present, than to any other source. This is so because intelligence is a value word that has been based on activities that are required for the continuance of western society. Such activities, in Africa, are learned to the point of stability almost wholly in schools. This sort of intelligence we recognise as a culturally-controlled derivative of Hebb's Intelligence B, and because its variance as measured by tests may not be static enough yet for precise apportionment to well-defined sources, it is termed Intelligence X (Irvine, 1963, 1964 pp. 364 *et seq.*). This meaning of intelligence is applied to all scores on tests applied in underdeveloped countries to groups for whom tests are unfamiliar and are given in English, a second language.

---

[1] It is hoped to present the details of these analyses in a future paper.

We come now to a general hypothesis for this type of study, carried out with groups who are literate in English, where wide differences in educational opportunity and culture-contact exist, and where environmental influences on test scores show patterns consistent with local conditions but are different from those generally found in western societies. This is :

As societies become more developed and industrialised, so their school systems become more homogeneous. Consequently (a) the effect of school quality on test performance will decrease in direct proportion to the increase in standards of teaching and school amenities and (b) the relationships of test scores to environmental factors will more resemble those of western societies.

This hypothesis accounts for the resemblances found between inter-correlation matrices that are derived from western groups and those from groups of non-western origin in underdeveloped countries and lays the blame for weaker relationships between test scores at the door of unpredictable group influences or censors that cut across individual differences. Our second factor analysis showed, for example, in a much more educationally homogeneous group, a more distinct set of group factors and this may indicate less censor variance due to inequalities in standards of education.

IV.—SUMMARY AND CONCLUSIONS.

With this general hypothesis it is perhaps appropriate to end the discussion and to summarise the main points. Educational measurement in Africa is complicated by unknown influences on test scores. Some of these are estimated to come from the following sources (Irvine, 1965).

(1) The content of the test itself.
(2) The form and style of the test.
(3) The transfer that takes place between practice items and actual test items especially when the material is unfamiliar.
(4) The particular cultural or educational bias of the test items.
(5) The motivational influence of strange testers who at present tend to be Europeans.
(6) Error variance.

Depending on the criterion used, these sources of variance could contribute to the specificity and unreliability of the test, so that constructors of tests for use in Africa must be in a position to limit as far as possible sources of variance that are irrelevant to prediction. Perhaps here is a fruitful area for programmed instruction, especially where trained test administrators are not available to carry through exacting test administration procedures such as at present appear to characterise the use of psychological tests in Africa (cf. Schwarz 1961).

There appears to be no reason why test constructors in western societies, moreover, should not experimentally determine the effect of different, controlled test presentations on test scores in order to find out at what point the test items become freely understood before the test itself begins, so that the score of even the dullest pupil is not contaminated by a failure to understand what is required by the test. Programmed instruction in pre-test procedures, which reduces the role of the administrator, may become a valuable tool for the tester. It certainly appears essential in Africa.

From an examination of these possible sources of variance, it was concluded that comparison of performance across cultures was a psychological task full of pitfalls. A technique for assessing the degree of comparison possible may include correlation of item difficulties, and it may be possible to use a frame of constructs devised for use in western societies to examine the contribution of such sources of variance to the meaning of test scores from particular groups.

The work in Rhodesia seems to indicate that constructs like ' g,' v.e.d. and ' n ' are valid for African schoolchildren at the end of the primary stage ; and that they can be used to choose psychologically meaningful variates for regression purposes. On the other hand, the influences of environment on scores at present would prevent the exact equation in meaning of scores on tests applied across cultures. In time it may be possible to bring such scores closer in meaning as societies become more homogeneous in educational opportunity and in the value they place on cognitive skills. The conditions to be met before comparisons can be made more accurately among underdeveloped and developed countries are that underdeveloped countries require many more skilled teachers to level up school differences, a large adult literacy programme to provide environmental supports for school learning and detailed studies of curriculum and methods related to child development so that teachers may be helped to teach more efficiently and with greater relevance to their school populations. The psychological field in Africa is large and it presents areas for study that may cause a closer examination of the general issues involved in the measurement of skills and abilities. In fact, one might predict that revaluations of psychological theories will come sooner if psychologists turn their energies to problems in underdeveloped countries.

V.—REFERENCES.

BIESHEUVEL, S. (1943). *African Intelligence.* Johannesburg : Inst. Race Relations.

BIESHEUVEL, S. (1952) Personnel selection tests for Africans. S.A *J. Science,* 49, 1, 3-12.

BIESHEUVEL, S (1954). The measurement of occupational aptitudes in a multi-racial society *Occup. Psychol.,* 28, pp. 4, ff.

BRIMBLE, A. R. (1963) The construction of a non-verbal intelligence test in Northern Rhodesia. *Rhodes-Livingstone J.,* 34, 23-35.

CENTRE D'ETUDE PROFESSIONELLE (1958). *Etude sur le P.M. de Raven.* Mimeographed Leopoldville.

FICK, M. L. (1939). *The Educability of the S African Native.* Pretoria . S.A. Council for Educ. Res.

IRVINE, S. H. (1963) Ability testing in English-speaking Africa. An overview of comparative and predictive studies *Rhodes-Livingstone J.,* 34, pp. 44-55.

IRVINE, S. H. (1964). *A Psychological Study of Selection Problems at the end of Primary Schooling in Southern Rhodesia.* Unpublished Ph D. Thesis, Univ. of London Library.

IRVINE, S. H (1965) Adapting tests to the cultural setting . A comment. *Occup. Psychol ,* 39, pp. 13-23.

KEIR, G. (1949). The progressive matrices as applied to school children. *Brit. J. Stat. Psychol.,* 2, pp. 140-150.

LORAM, C. T. (1917). *The Education of the South African Native.* London : Longmans.

MACARTHUR, R. S., IRVINE, S. H., and BRIMBLE, A. R. (1964). *The Northern Rhodesia Mental Ability Survey,* 1963 *Rhodes-Livingstone Institute, Lusaka.*

MACDONALD, Lt.-Col. A. (1944-5). *Selection of African Personvel.* Reports on the work of the Selection of Personnel, Technical and Research Unit, M.E.F. War Office Archives .London.

MORRISBY, J. R. (1955). *Manual for the Compound Series Test.* London . N F.E.R.

NISSEN, H. W. MACHOVER, S., and KINDER, E. F. (1935). A study of performance tests given to a group of native African Negro children. *Brit. J Psychol.*, 25, pp. 308-355

OLIVER, R. A. C. (1932). The comparison of the abilities of races with special reference to East Africa. *East African Medical J.*, 9, pp. 160-204.

OLIVER, R. A. C. (1933). Adaptation of intelligence tests to tropical Africa. *Overseas Educ.*, 4 (4) and 5 (1).

OLIVER, R. A. C. (1934). Mental tests in the study of the African. *Africa*, 7, pp. 40-46.

SCHWARZ, P. (1961). *Aptitude Tests for Use in the Developing Nations.* Pittsburgh : Amer. Inst. Res.

SILVEY, J. (1963). Aptitude testing and educational selection in African. *Rhodes-Livingstone J.*, 34, pp 9-21.

SLATER, P. (1948). Comment on " The comparative assessment of intellectual ability." *Brit. J. Psychol.*, 39, 20-21.

VERNON, P. E. (1950). An application of factor analysis to the study of test items. *Brit. J. Stat Psychol*, 3, 1-15.

VERNON, P. E (1958). *Educational Testing and Test-Form Factors.* Princeton N.J. E.T.S. Res. Bull. 58-3.

VERNON, P. E. (1961a). *The Structure of Human Abilities.* London : Methuen.

VERNON, P. E. (1961b). *Selection for Secondary Education in Jamaica.* Kingston : Govt. Printer.

(*Manuscript received* 21*st January*, 1965)

VERBALISATION EFFECTS IN A 'NON-VERBAL' INTELLIGENCE TEST

By J. RADFORD*†

(*Birkbeck College, University of London*)

SUMMARY. Two experiments were carried out with a total of ninety children aged between 9 and 12 to investigate the effects of verbalising on problem solving. The problems were those of the Progressive Matrices intelligence tests. It was found that performance improved when subjects were asked to give reasons for their choices and these reasons were verbally reinforced. Reinforcement is thought to have had the effect of decreasing stereotyped responses. An analogy between intelligence tests of this type and Harlow's (1949, 1959) 'learning sets,' suggested by Cashdan, 1961, is rejected.

I—INTRODUCTION.

THE work of Luria, 1961, and others, suggests that ability to verbalise is an important factor in problem solving and concept attainment. The abilities involved in these activities appear to be measured by some intelligence tests, which therefore form suitable instruments for their investigation. In an experiment by Cashdan, 1961, it was hypothesised that if children carrying out the Progressive Matrices (children's version) test of intelligence were told to give reasons for their choices, their performance would improve. This effect was not found. A different approach is suggested by work on verbal conditioning. Many experimenters have shown (Salzinger, 1959) that verbal responses can be significantly modified by suitable reinforcement, itself usually verbal. The total number of words spoken, and/or the frequency of specific words or phrases, may be controlled. Further, Cashdan suggested that performance on a test such as the Matrices can be regarded as analogous to the formation of 'learning sets.' This term was introduced by Harlow in 1949, to refer to the behaviour of animals trained on a series of problems of similar type and equal difficulty. Typically, subjects (rhesus monkeys) became able, after such training, to solve problems that were previously beyond their capacities. A number of experimenters have extended these methods to children, of both normal and subnormal intelligence (Shepard, 1957 ; Kaufman and Peterson, 1956 ; House and Zeaman, 1959). Learning set procedure involves the giving of knowledge of results, which, of course, is not standard practice in the Matrices test, but if this is done it would appear reasonable to expect somewhat comparable effects.

With these findings in mind two experiments were carried out, designed to investigate performance on the problems of the Progressive Matrices tests under various conditions of verbalisation and reinforcement and with opportunity for more than one trial per problem.

II.—METHOD.

(1) *Experiment* 1.

(i) *Subjects.* Subjects were eighteen children, seven girls and eleven boys, aged 9–10, from a junior modern school, and had a mean I.Q. of 110·4, range 103-121, on the NFER Intelligence Test III.

---

* Now at West Ham College of Technology.

† This work was carried out under a Medical Research Council Studentship for Training in Research Methods, held from 1961 to 1964. It forms part of a thesis accepted by the University of London for the Ph.D. degree.

(ii) *Material.* Set C of the Progressive Matrices (1938) was selected as being likely to be of sufficient difficulty for these subjects under the experimental conditions. Each page of the test contains a 3×3 matrix, with the last item missing. The principle embodied in each matrix is expressed in both columns and rows. Thus, in theory, any item of the nine may be missing without altering the difficulty of the problem. Nine versions of the twelve problems of Set C were prepared, each with a different item missing. The display from which S was required to select the correct item comprised twelve items—the nine of the problem, plus three irrelevant ones. Positions of the various items were randomised.

(iii) *Procedure.* The problems were given individually with the normal Matrices procedure, except that subjects were randomly placed into one of three groups : (1) the subject was given knowledge of results after each trial ; (2) the subject was, in addition, asked to give the reason for his choice, which was received by E in silence ; (3) in addition to the foregoing, E attempted to reinforce selectively the more appropriate responses, by saying ' good,' ' very good,' and/or ' yes.'

In order to have a further test of the effects of verbalisation, all Ss, after Set C, completed Set D of the Matrices in the normal way, that is by writing down the answers, without giving reasons.

(2) *Experiment* 2.

(i) *Subjects.* Subjects were seventy-two children aged 11–12 from a secondary modern school. They were allocated to six matched groups on the basis of scores on the N.F.E.R. Intelligence Test III, scores ranging from 75–104, Mean 88·6. Each group comprised seven boys and five girls.

(ii) *Procedure.* Each subject was given individually the Progressive Matrices (children's coloured version) test of intelligence. Contrary to the standard procedure, after each response, the subject was told whether or not he was correct. If incorrect, two further attempts were allowed. This constituted a partial correction procedure, and was adopted in order to control for the fact that subjects tended to embody a correction in their verbal reasoning. The first response was taken for scoring purposes, however, so as to avoid a spurious superiority of the verbal conditions. Three of the groups were asked after each response to say why they had made that choice. In addition, three conditions of reinforcement were used : one in which the experimenter received the subject's response in silence ; one in which he replied with ' good ' (or ' good try ' in the case of incorrect responses) to every response ; and one in which an attempt was made to reinforce only appropriate verbalisations, that is, those that showed some grasp of the principle involved in the particular problem. This was intended to ' shape ' the behaviour towards more precise reasoning. In the non-verbal conditions reinforcement followed the choice. All responses were recorded on tape.

III.—RESULTS.

(1) *Experiment* 1.

Table 1 shows the average numbers of correct responses per problem for the three groups. The groups differ significantly at the ·05 level, and problems at the ·001 level.

TABLE 1

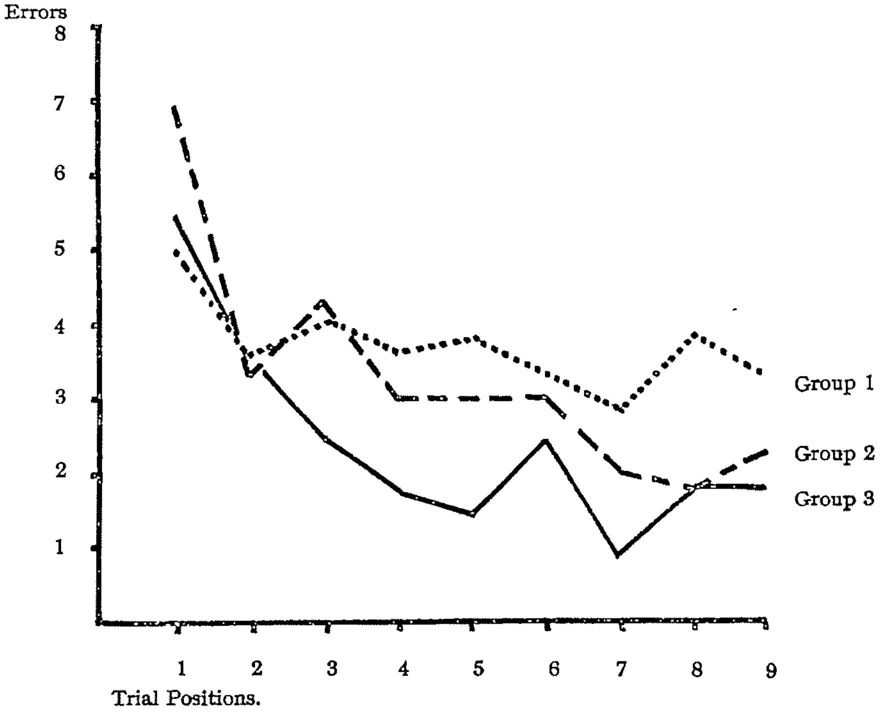
AVERAGE CORRECT RESPONSES OVER THE COURSE OF TWELVE PROBLEMS IN EXPERIMENT 1 (total possible score per problem=9).

Problem	1	2	3	4	5	6	7	8	9	10	11	12
Group 1	4·0	9·0	6·0	5·4	5·7	7·4	5·0	7·4	4·7	6·0	5·4	6·5
Group 2	5·9	7·0	7·0	5·2	7·5	8·2	5·2	7·6	6·0	5·7	6·7	5·7
Group 3	5·2	7·8	6·7	7·7	7·2	9·0	5·7	8·7	6·5	7·0	6·2	6·5

Groups differ significantly at the ·05 level, and problems at the ·001 level.

Fig. 1 shows the nature of this difference : the tendency is for Group 3 to draw away from the other two groups after Trial 2 or 3. That is, these Ss are quicker to solve successive problems. Similarly, Group 2 is superior to Group 1. No significant differences were found between the groups on Set D.

FIG. 1. TRIAL POSITIONS IN EXPERIMENT 1.



Trends : Group 1 not significant at ·05 level.
Group 2 significant at ·01 level.
Group 3 significant at ·05 level.

(2) *Experiment* 2.

Table 2 gives the average scores for the six conditions on the three sections of the test. Analysis of variance shows that subject differences are significant throughout (·05 level in Sections A and B, ·01 level in Section Ab). On the first two sections the variance due to verbalisation is not significant, but in Section B it is significant at the ·01 level. No other significant differences appear, i.e., from the reinforcement variable or from the interactions. Comparison of the means for the six conditions shows that in Section B the condition of verbalisation with appropriate reinforcement (Condition 6) is significantly better than all three non-verbal conditions (·05 level), and particularly than the non-reward, non-verbal condition (·01 level). No other significant differences appeared between means.

TABLE 2

AVERAGE CORRECT RESPONSES UNDER SIX CONDITIONS IN EXPERIMENT 2 (total possible score per section = 12).

Condition	1	2	3	4	5	6
Section A	9·58	9·33	8·92	10·00	9·50	10·17
Section Ab	8·83	8·92	8·58	9·42	9·58	9·17
Section B	7·33	7·67	7·67	8·75	8·83	9·25

Subject differences significant at ·05 level in Sections A and B, at ·01 level in Section Ab. Verbalisation difference significant at ·01 level in Section B

The average numbers of words spoken by subjects in the three verbalisation conditions were examined similarly. No significant differences appeared here. Thus, it appears from both experiments that reinforced verbalising results in improved performance on a ' non-verbal ' test.

Two other aspects of the results should be noted. First, a number of investigators, e.g., Koch and Meyer, 1959, have reported a relation between intelligence and the formation of learning sets. The latter is thought to involve abilities other than those of simple learning, which have often been reported (e.g., by MacKay and Vernon, 1963) not to correlate with intelligence. In the present case, it would be expected that those subjects scoring more highly on the N.F.E.R. test would show greater gains on the experimental task (as distinct from merely scoring more highly throughout). Rank correlations were calculated between N.F.E.R. scores and the differences between scores on Sections A and B of the Matrices, as an index of transfer. It was found that none of the correlations was significant at the ·05 level, but those for Conditions 4, 5 and 6 are in the predicted direction.

Lastly, since the verbal conditions appear to be distinguished from the non-verbal, it is interesting to examine the content of the protocols here. The most obvious point here is that there is a marked difference between the three conditions in respect of the number of steriotyped responses. Many of the subjects produced runs of identically worded responses. Table 3 shows these. It is clear that they are more frequent in Conditions 4 and 5, and in the less intelligent subjects.

TABLE 3

RUNS OF THREE OR MORE IDENTICAL RESPONSES IN VERBAL CONDITIONS IN EXPERIMENT 2.

S	Condition 4			Condition 5			Condition 6		
	A	Ab	B	A	Ab	B	A	Ab	B
1.	10	6, 3	5, 6	3	3				
2.			3, 3	3, 4	11				
3.	5	3		5	8, 5	10			
4.								3	3, 3
5.	12	11	12	9	3				
6.	3	3		5	3	3			
7.				3			11	7	
8.	3	3		3					
9.									
10.									
11.	3, 3	3							
12.		3							

Each separate run of three or more identical responses is entered. In each condition subjects are in matched order in respect of N.F.E.R. scores, 1. being the lowest, and 12. the highest.

IV.—DISCUSSION.

The experiments show that appropriately rewarded verbalising produces higher scores on problems of the Matrices type, and this effect is progressive. That is to say, subjects between the various conditions become more widely differentiated over the series of problems. So also do subjects within conditions; the more intelligent subjects gain more from the experimental treatment. This gain shows itself verbally in a decrease in stereotyped responses, as well as in higher scores. Although, as Cashdan points out, there is an apparent similarity to a learning set situation, it is felt that this is superficial. It is argued that what is learnt in acquiring a learning set—what is transferred from one problem to the next—is a strategy or operation equally applicable to all problems of a series by virtue of their common structure. In the present case such a strategy cannot be distinguished. Rather, what seems to be acquired is a greater awareness of the sort of principles involved, which are not necessarily the same from one problem to the next, and a better cognitive set towards the task, somewhat after the pattern suggested by Fitts, 1964. The findings also illustrate some of the general principles derived from work on operant conditioning which are of the greatest importance for learning: the effect of knowledge of results; the adequate delimitation of step size; immediacy of reinforcement; and active *versus* passive response. These principles operate here to modify the mediating (verbal) processes involved in problem solving, making them progressively more appropriate and thus resulting in more frequent and/or more rapid solution.

V.—REFERENCES.

CASHDAN, A. (1961). *Problem-Solving and Learning Sets in Normals and Subnormals.* Proceedings of the London Conference of the British Psychological Society, Dec., 1961 (abstract).

FITTS, P. (1964). Perceptual-motor learning. In Melton, A. W. (ed.) *Categories of Human Learning.* New York: Academic Press.

HARLOW, H. F. (1949). The formation of learning sets. *Psychol. Rev.*, 56, 51-65.

HARLOW, H. F. (1959). Learning set and error factor theory. In Koch, S. (ed.) *Psychology A Study of a Science*, Vol. 2. New York : McGraw-Hill.

HOUSE, B. J., and ZEAMAN, D. (1959). Position discrimination and reversals in low-grade retardates. *J. Comp. Physiol. Psychol.*, 52, 564-565.

KAUFMAN, M. E., and PETERSON, W. M. (1958). Acquisition of a learning set by normal and mentally retarded children. *J. Comp. Physiol. Psychol.*, 51, 619-621.

LURIA, A. R. (1961). *The Role of Speech in the Regulation of Normal and Abnormal Behaviour* (ed. Tizard, J.). Oxford : Pergamon.

MACKAY, G. W. S., and VERNON, P. E. (1963). The measurement of learning ability. *Brit. J. Educ. Psychol*, 33, 177-186.

SALZINGER, K. (1959). Experimental manipulation of verbal behaviour : a review. *J. Genet. Psychol.*, 61, 65-94.

SHEPARD, W. O. (1957). Learning set in pre-school children. *J. Comp. Physiol. Psychol.*, 50, 15-17.

(*Manuscript received* 29*th March*, 1965)

THE HAPPINESS OF PUPILS IN CO-EDUCATIONAL AND SINGLE-SEX GRAMMAR SCHOOLS: A COMPARATIVE ASSESSMENT

By R. R. DALE

(*Senior Lecturer in Education, University College of Swansea*)

SUMMARY. A questionnaire was administered to some 1,500 ex-pupils of grammar schools, who were students in teacher training courses. It included questions on the pleasantness of the school atmosphere, the happiness of life in school, and the preference for mixed or single-sex schools. The sample is admittedly not representative of all types of ex-pupils. The differences found between the co-educated women and women from girls' schools were highly significant with regard to the pleasantness of the school atmosphere and the happiness of life in school, whereas the differences between the men's groups were not significant. Men and women educated in mixed schools were almost unanimous in their favour, whereas those from single-sex schools were very divided.

I.—INTRODUCTION

RESEARCH on the value of co-education as compared with education in single-sex schools, at the secondary level, has until recently been concentrated on academic attainment in the grammar school. It has now moved to the emotional and social aspects, on which very little work has been done apart from a few researches concerning the attitude of pupils towards school subjects, and only part of this is relevant. Some years ago Lotz (1952), in New York, compared the emotional and social development of boys and girls educated in a co-educational high school with that of pupils in two single-sex schools, testing some pupils at entrance and comparing their scores with those of groups of older pupils in the same schools. He concluded that the type of schooling made little difference to the boys, but that the co-educational schooing was better for the girls. As, however, his sample consisted of only three schools, the findings cannot safely be generalised, especially to British institutions. In Britain Moreton (1946), Davies (1950) and Dale (1955 and 1965) have reported their findings in the related field of teacher opinion on co-education. These indicated that opinion among secondary grammar teachers is in favour of co-educational secondary schools, and demonstrated the influence of the teacher's own education on his or her preference, though this was much stronger in the case of the co-educated than the single-sex educated. Dale (1948 and 1949) has also published two slight preliminary articles on the attitudes of sixth formers and of small samples of grammar school pupils of other ages about co-education. All these groups also preferred co-educational secondary schools, though not all the sub-groups had direct experience of co-educational schooling. The present enquiry was designed to make an exploratory survey into the factors associated with this preference for co-education. The writer then intended to make a more detailed investigation wherever the survey indicated that this was most likely to give significant findings.

II.—PROCEDURE

For this exploratory work the questionnaire approach was considered suitable. It quickly covered a broad area, was simple to understand and answer, and its anonymity was a decided advantage. As exactly the same questions were asked of co-educated and single-sex educated men and women, usually without the type of school being mentioned, the possibility of bias in the questions was reduced. Opportunity for comment was provided as a supplementary to a number of the questions, in order to discover the reasons for the replies, to detect possible ambiguities in the wording of the question, and to discover whether respondents wished to qualify any of their answers.

The questionnaire was administered to over 1,500 intending teachers who were in training colleges and university departments of education. The proportion of students from mixed and single-sex training colleges was 6 to 5. Almost half were women. Respondents were asked to complete the questionnaire as ex-pupils of grammar and comprehensive schools. They are, therefore, limited to one type of ex-pupil and are not respresentative of such schools as a whole. On the other hand they are a fairly homogeneous body of students who intend to make teaching their profession and are likely, therefore, to have a genuine interest in educational problems and a responsible attitude to research. No attempt was made to obtain representative samples of intending teachers but the percentage of training college students at 77·4 is only slightly smaller than the 85 per cent. in the whole population of student teachers, taking as a basis the number of trainees entering the profession each year. Adjusting the proportion would have made only a very slight difference to the results (in favour of co-education); as other factors would also have had to be taken into account in any attempt to make the sample representative, it was decided to present the results as they stood.

The questionnaire was administered under examination conditions to large groups of students, seven groups by the writer and the remainder by college lecturers. The percentage of completed forms was, therefore, 100 per cent. Though most of the forms were completed by first year training college students in their first term, there was no attempt made to exclude the influence of the college course itself. The results are, however, in line with those reported, on a small scale, from sixth form pupils of grammar schools. (Dale, 1949.)

For most of the questions the ex-pupils were asked to indicate their value-judgment on a five-point scale, ranging for example from ' Like very much ' to ' Dislike very much,' by placing a tick against the value they chose. The percentage of ex-pupils who selected each value was then calculated, keeping the sexes separate and comparing the judgments of the co-educated with those of ex-pupils of single-sex schools. The differences were assessed for statistical significance by the chi-square test. Though there were a large number of questions lack of space limits the article to the results from three of them. First came, ' Was your life in school, viewed as a whole : Very happy, happy, jogging along, rather unhappy, very unhappy ? ' A space for comment was left after the question. The second question, meant to link with the first, and set in part as a check on it, was, ' Do you think the school atmosphere was, Very pleasant, Pleasant, Neutral, Rather unpleasant, Very unpleasant ? ' There was no provision for comment.

The third question had two versions : ' If you attended a single-sex school would you have preferred it to have been a mixed school ? ' and, ' If you attended a mixed school would you have preferred it to have been single-sex ? ' Shortage of time precluded opportunity for comment.

III.—RESULTS.

(*a*) *Happiness in school.*

The replies to the first question are presented in Table 1.

TABLE 1

Ex-pupils of....	Men		Women	
	Single-sex schools	Mixed schools	Single-sex schools	Mixed schools
Very happy	149 (30·2%)	101 (34·2%)	109 (24·5%)	121 (38·5%)
Happy	238 (48·3%)	146 (49·6%)	247 (55·7%)	162 (51·6%)
Jogging along	91 (18·5%)	40 (13·5%)	79 (17·8%)	27 (8·6%)
Rather unhappy	14 (2·8%)	7 (2·4%)	8 (1·8%)	4 (1·3%)
Very unhappy ..	1 (0·2%)	1 (0·3%)	1 (0·2%)	0
Totals	493 (100%)	295 (100%)	444 (100%)	314 (100%)

The most noticeable feature is that few of these ex-pupils (who would have been at least fairly successful in school work) were rather unhappy or worse. On inspection there is no practical or statistically significant difference between the proportion of these unhappy individuals in mixed schools and that in single-sex schools. On the positive side, however, there is a decided difference between the co-educated and single-sex educated women, the former being more positively happy when at school than the latter. This greater happiness is to be seen in the 14 per cent. superiority in the ' Very happy ' section and in the 8 per cent. fewer who admitted to have been only ' Jogging along.'* The 4 per cent. superiority of ex-pupils of girls' schools in the ' Happy ' category redresses the balance only slightly. Though the trend for the men is the same as for the women, the difference between the two groups is not statistically significant.

Slightly less than half the respondents wrote a comment. Only with the ex-pupils of boys' schools did the proportion rise slightly above one-half. The answers revealed one difficulty in the answering of the question, namely that a few ex-pupils had trouble in assessing their happiness throughout their secondary school life because it either fluctuated so much, or was very different in the lower school compared with the upper school. Of the 115 ' old girls ' of mixed schools who made comments, two said their happiness varied very much, six said the upper school was happier and four said they were not as happy in the sixth : one was happier in the lower school. Of the 182 ex-pupils of girls' schools who commented there were eight who found the lower school happier, and fifteen who were discontented with their sixth form life ; on the other hand there were thirty-two who were happier in the upper school or sixth form. Only a small proportion of these said they had difficulty in making an overall evaluation and none of them failed in fact to express a judgment. The full classification is given in Table 2.

Though there were over 1,500 completed forms and nearly 700 of these included comments (about 800 when classified, because some included two aspects), when these were divided among the four main groups and classified into some ten categories of comment, the resulting numbers were inevitably small. The Table is, therefore, published not as something conclusive, but as a

---

* The difference between the overall distributions in mixed and girls' schools is statistically highly significant (beyond the ·001 level).

TABLE 2

CLASSIFIED COMMENTS ON HAPPINESS IN SCHOOL.

	Female Students		Male Students	
School Education	Girls' Schools	Mixed Schools	Boys' Schools	Mixed Schools
Pupil/staff relations happy	22	35	39	25
Pupil/staff relations unsatisfactory	8	6	19	8
Atmosphere good	21	26	15	20
Atmosphere unpleasant	2		1	1
Discipline good	2	1		
Discipline too strict	11	1	3	
Societies and sports enjoyable	10	15	31	17
Societies too few	2		5	3
Friendships	27	23	27	21
Work enjoyable	10	16	19	10
Work difficulties*	14	7	22	13
Happiness varied	3	2	5	1
Lower school happier	8	2	2	
Upper school happier	32	5	30	11
Sixth Form discontent	15	4	4	
Mixed school happy		4		4
Personal difficulties		2	6	
Others	23	15	33	24
Totals (of comments)	210	164	261	158

* 'Work difficulties' includes finding school work not interesting or difficult, the pressure of examinations and excessive homework.

tentative qualitative illustration of the quantitative part of the previous Table. Further work is in progress which may give more definite results. Here one can only point out interesting trends, keeping in mind that the subjects are confined to one type—intending teachers.

If we compare the two groups of female students we find a few of these trends. Those from mixed schools appear, on the whole, to have experienced rather happier relations between pupils and staff than did those who were in girls' schools—though very few of either group said that those relations were unsatisfactory. There may be some measure of support for this in the comments on the school atmosphere and school discipline, though there may be two opposing views about such a thorny question as discipline! Perhaps the presence of men on the staffs of mixed schools may have influenced these replies, in which case the situation may now be changing somewhat, as more and more girls' schools are appointing a small proportion of men to their staffs. Whatever the reasons, we have to record that, whereas four in ten of the comments from women ex-pupils of mixed schools were favourable,* the proportion from girls' schools was little more than two in ten; similarly, whereas less than 7 per cent. of the comments from the female ex-pupils of mixed schools were adverse, the percentage was 17 from girls' schools. The mixed schools also have some advantage in the comments on work and in the relative lack of sixth form discontent. It might seem that the girls' schools have an advantage in that a much greater proportion of their ex-pupils say they were 'happier in the upper school,' but this is merely in apposition to their less happy experience in the lower school; the ex-pupils of mixed schools appear only rarely to experience this variability, or find the change so slight that they are not impelled to comment.

When we compare the two groups of male ex-pupils the most marked difference is in the greater variability of the happiness of boys in boys' schools. One-seventh of them were happier in the upper than the lower school; in the mixed schools this proportion was less than one-fourteenth. If the categories of pupil-staff relations, atmosphere, discipline and 'mixed school happy' are combined there is a trend in favour of the mixed school, in line with the trend shown by Table 1. Almost one-third of the boys from mixed schools comment favourably, but only one-fifth of those from boys' schools; again, while about one in eighteen from the mixed schools comment adversely in these sections, the proportion from the boys' schools is one in ten.

(*b*) *School atmosphere.*

The replies to the second question are given in Table 3.

TABLE 3

School Education	Men		Women	
	Single-sex	Mixed	Single-sex	Mixed
Very pleasant	106 (21·4%)	79 (26·8%)	61 (13·6%)	103 (33·0%)
Pleasant	273 (55·1%)	147 (49·8%)	264 (58·9%)	171 (54·8%)
Neutral	91 (18·3%)	58 (19·7%)	98 (21·8)%	34 (10·9%)
Rather unpleasant	25 (5·0%)	10 (3·4%)	24 (5·4%)	4 (1·3%)
Very unpleasant ..	1 (0·2%)	1 (0·3%)	1 (0·2%)	0
Totals	496 (100%)	295 (100%)	448 (100%)	312 (100%)

As was the case with the first question, few pupils found the school atmosphere to be rather unpleasant or worse. Here only the women revealed a difference of practical significance between the two groups.† Just over one in a hundred of the co-educated women judged the school atmosphere to be 'rather unpleasant,' whereas the proportion for women from girls' schools was rather more than one in twenty.† Once again, however, it is on the positive side that the chief difference lies, and likewise it is only with the women that the difference is large and statistically significant. *Whereas one-third of the girls from mixed schools assessed the school atmosphere as 'Very pleasant,' not quite one-seventh of the girls from girls' schools were as enthusiastic. Whereas a ninth of the co-educated girls evaluated the atmosphere as 'neutral,' more than a fifth of the opposite group registered the same assessment. It must be understood that one is here merely reporting results and that these can have various interpretations. Those women who are teaching in girls' schools and are devoted to their profession may understandably feel hurt and misjudged when they read these figures, but unnecessarily so for several reasons. First, because research may reveal other aspects in which girls' schools are superior; second, because we are dealing with averages and there are certainly good girls' schools and poor mixed schools as well as *vice versa*; third, because the nature of the results may be due mostly to the simple fact that a mixed staff and mixed pupils in themselves produce a better atmosphere—and it may be that no woman teaching in a girls' school could do anything to prevent this difference arising.

---

* Combining 'Happy staff-pupil relationships,' 'Good atmosphere,' 'Good discipline' and 'Mixed school happy.' The adverse comments combine the reverse of the above, together with 'Sixth Form discontent.'

When these adverse and favourable comments for mixed and single-sex schools are opposed in a 2×2 table, the resulting Chi-square is significant well beyond the ·01 level for girls and beyond the ·025 level for boys.

† A chi-square test shows the difference to be significant well beyond the ·01 level in a 2×2 table, with unpleasant categories combined and opposed to the three other categories.

(c) *School preference.*

The third question asked whether the respondents would have preferred a mixed or single-sex school. The precise wording is given with the Tables. The preferences expressed give strong though indirect support to the results from the two previous questions. Table 4 shows that the men who were educated in boys' grammar schools were uncertain whether they would have preferred a mixed or single-sex school, whereas a majority (53 per cent.) of the ex-pupils of girls' schools (with a fifth undecided) would have preferred a mixed school.

TABLE 4

'IF YOU ATTENDED A SINGLE-SEX SCHOOL WOULD YOU HAVE PREFERRED IT TO HAVE BEEN A MIXED SCHOOL?'

	Male Ex-pupils	Female Ex-pupils
Yes (preferred mixed school)	189 (39·8%)	235 (53·2%)
No (preferred single-sex school)	196 (41·3%)	110 (24·9%)
Undecided	90 (18·9%)	97 (21·9%)
	475 (100%)	440 (100%)

The verdict of the former pupils of mixed schools is much more dramatic, as seen in Table 5.

In great contrast to the indecision of the boys from boys' schools, and to some extent the girls from girls' schools, both boys and girls from co-educational schools are very decided in their preference for a mixed school. As many as 90 per cent. of the boys and almost 96 per cent. of the girls would not have preferred their school to have been single-sex; only twelve boys in 293 and five girls in 307 wished that their mixed school had been a single-sex one. Human nature and human temperaments being so varied one could scarcely expect to get a smaller number in the minority no matter what the question asked!

*Later figures for men increased the difference: p becomes greater than 0·025

TABLE 5

' IF YOU ATTENDED A MIXED SCHOOL WOULD YOU HAVE PREFERRED IT TO HAVE BEEN SINGLE-SEX ? '

	Male Ex-pupils	Female Ex-pupils
Yes (preferred single-sex school)	12 (4·1%)	5 (1·6%)
No (preferred mixed school)	264 (90·1%)	294 (95·8%)
Undecided	17 (5·8%)	8 (2·6%)
	293 (100%)	307 (100%)

Unfortunately, as there was no space left for comments, the reasons for this preference cannot be given, except in so far as these have been expressed in answer to the other questions. Only two of these concern us here, those concerning happiness in school and school atmosphere. That the girls in mixed schools evaluate themselves as more happy than do those in girls' schools, and also place a higher evaluation on the school atmosphere, in part explains their heavy preference for a mixed school. On the other hand the differences between the two boys' groups are small and cannot account for the strength with which the boys from mixed schools express their preference for mixed schools. For an analysis of their reasons we shall have to wait for the findings of further research.

IV.—DISCUSSION.

The results from the first two questions are very similar, and each helps to confirm the other.* The fact that the question on school atmosphere throws up a rather greater difference than that on ' happiness in school ' is interesting. One might well have expected the reverse because some of the single-sex schools have an atmosphere steeped in tradition, which might have increased the pupils' evaluation of the ' school atmosphere.' If this factor exists it appears to be heavily outweighed by the regard of the co-educated women for what protagonists of mixed schools call the ' natural family atmosphere ' and the ' natural relationship between the two sexes in work and play.' Nor does it result in a higher evaluation of the ' atmosphere ' by boys from boys' schools, compared with those from mixed schools.

Unfortunately, we do not know the exact number of schools involved in the present survey, but it is likely to be a minimum of one hundred in each group. The training colleges and university departments of education taking part were in many districts of England and Wales, and each of these institutions would take pupils from a large number of schools, thus making the research broadly based. On the other hand there would probably be some under-representation of ex public school pupils.

In this kind of work one is always faced with the problem of whether the evaluations expressed were those which were really felt. Ultimately, there is no adequate proof that this is so unless corroborative objective criteria can be produced, and this is far from easy to obtain, though we do hope to be able to provide some in time. Meanwhile we can only say, first, that by making the replies anonymous we endeavoured to remove both the desire to please and the fear of being thought disloyal; second, that the questionnaire and its administration were exactly the same for ex-pupils from both co-educational and single-sex schools, and we ourselves cannot see any reason why pupils from one of these groups rather than the other should have failed to put down what they really thought. Indirect support for the general validity of these findings is, however, to be found in the results of other related research such as those quoted earlier, by Moreton, Davies, and Dale, which demonstrated the preference of teachers for mixed secondary schools. If the teachers themselves were happier in mixed schools it is likely that the pupils would also be happier.

---

* The closeness of association is demonstrated by a contingency co-efficient of 0·63 for girls from girls' schools, and 0·50 for girls from mixed schools. When corrected for fineness of grouping (in a 5×5 Table) these became 0·71 and 0·56, respectively, and (by conversion to Chi-square) are highly significant. It is not, however, contended that the distributions should be exactly the same for the replies to the two questions. The coefficients for the boys were not calculated.

Another problem which arose is of particular interest as it illustrates the hidden dangers which are inherent in these investigations. There tends to be a difference in social class composition between single-sex and mixed grammar schools, deriving in part from the urban location of the one and the rural location of the other, and also from such factors as the age and independence of the school. The writer thought it possible that there might be a variation in the evaluation of 'happiness in school,' etc., correlated with social class, thereby creating a systematic hidden factor which would lead to a wrong interpretation of the results. Fortunately, the necessary information about the social class of a large section of the respondents was readily available and a check was made, taking the sexes separately and also the two types of school. Only a small difference was, in fact, found between the social class composition of the boys in mixed and single-sex schools. This was in the expected direction, and its smallness might be because we are concerned, not with an overall representative sample, but only with intending teachers and their parents. Strangely, the difference was greater between the girls' schools and girls from mixed schools. The girls' schools had 39 per cent. of the sample in a broadened Class I, the mixed schools having only 26 per cent., though they had a 5 per cent. compensation in Class II; at the other end, the girls' schools only had 5 per cent. in Class IV and V, whereas the mixed schools had 11 per cent. If the analysis of the answers by parental occupation had now shown a correlation between occupation and the pupils' answers on happiness and school atmosphere the interpretation of the results would have been very different. But no such trend emerged, the scores for each category being remarkably similar, with two exceptions. In the girls' schools the pupils whose fathers were in Occupation Class I were the least happy, while in the mixed schools the least happy group had fathers in Occupational Class III. As there were only some 600 in the sample the sub-sections were small, and one would not rule out entirely the possibility that with a larger number of cases some very low correlation between the father's occupational class and the pupils' evaluation of the school atmosphere, etc., might emerge. The evidence from this study indicates, however, that the pupils' evaluations were not correlated with occupational class.

A sound interpretation of the evidence, therefore, seems to be that girls who later became teachers found life in mixed schools appreciably happier than similar girls found life in girls' schools. Though the trend for boys (who became teachers) is in the same direction it is small and statistically insignificant. Questions about 'the school atmosphere' have a similar result.* It is tempting to jump to the conclusion that the difference between the evaluations is due to the single-sex schooling on the one hand and the co-educational schooling on the other. This, indeed, would appear to be a likely answer. Before committing oneself to this, however, one would like to examine the effect of the school setting. Many more of the co-educational schools than the single-sex have a small-town or rural setting, and this contrast in environment may have an unconscious effect on the evaluation—unconscious because this was not mentioned in a single comment. Most of the comments are concerned with relations with people—friendship, teacher-pupil relations, general attitude of other pupils, etc. In the writers judgment the nature of the setting will prove to be less important than whether the schooling is mixed or single-sex, but we need more evidence before we can be more definite.

---

*The difference in favour of co-educated men has since become significant with a larger sample.

It might be objected that schools are intended for the promotion of attainment and the teaching of skills, and that the quality or depth of pupils' happiness is merely incidental. This is a narrow outlook and it ignores the likelihood that, in the long term, happy people are likely to work better. Most educationists, if not all, would also agree that education should seek to assist the emotional and social development of the individual. At one time some opponents of co-education would have postulated that co-educational pupils are happier because they don't work as hard as pupils in single-sex schools. Such an argument is scarcely tenable in the light of the critical analyses of research evidence on this subject which have recently been published. (Dale, 1962 and 1964.)

The situation described is not necessarily static. It could be affected by a change in public opinion and by changes in the schools themselves, such as the rapidly increasing percentage of married women and of men on the staffs of girls' schools.

ACKNOWLEDGMENTS.—The writer acknowledges with gratitude the financial aid given by the Department of Education and Science for his research in this field. Thanks are also expressed to those Professors and Lecturers of Education and Training College staffs who gave assistance. Dr. Phillip Williams is thanked for kindly reading the manuscript and making some useful suggestions, and Mrs. R. Lewis for her very efficient secretarial assistance.

V.—REFERENCES.

DALE, R. R. (1955). Co-education: the verdict of experience. *Brit. J. Educ. Psychol.*, 25, 3.

DALE, R. R. (1965). Co-education: the verdict of experience, II—A qualitative approach. *Brit. J. Educ. Psychol* (in the press).

DALE, R. R. (1948). Preliminary article. *Times Educ. Supp.* 28th August.

DALE, R. R. (1949). Preliminary article. *A.M.A.* September.

DALE, R. R. (1962). Co-education —I: A critical analysis of research on the effects of co-education on academic attainment in grammar schools. *Educ. Res*, IV, 3.

DALE, R. R. (1962) Co-education—II: An analysis of research on comparative attainment in mathematics in single-sex and co-educational maintained grammar schools. *Educ. Res.*, V, 1.

DALE, R. R. (1964). Co-education—III: Research on comparative attainment in English in single-sex and co-educational grammar schools. *Educ. Res.* VI, 3.

DAVIES, W. E. (1950). *A Study of the Attitudes of Secondary School Teachers towards Co-education, Differentiation of Curriculum and Sex Teaching.* Ph.D. thesis, London.

LOTZ, H. V. (1952). *The Relationship between Emotional and Social Adjustment of Individuals and their Attendance at Co-educational and Single-sex High Schools.* Ph.D. thesis., Sch. Educ. New York Univ.

MORETON, F. E. (1939) *Co-education—A Statistical Enquiry into the Attitude of Teachers towards Co-education, and a Comparative Study of the Emotional Development of Children Trained in Co-educational and Other Institutions.* Ph.D. thesis, University of London Institute of Education. (Summarised in *The Brit. J. Educ. Psychol*, June, 1946.)

(*Manuscript received 8th April*, 1965)

THE PREDICTION OF VOCATIONAL SUITABILITY FROM SECONDARY MODERN SCHOOL RECORD CARDS

BY JOHN H. EYRE

(*Headmaster, Lowfield County Secondary School, York*)

SUMMARY. Quantitative assessments of the school, employment and further education records of over 400 secondary modern school boys and girls were intercorrelated and subjected to factor analysis. Significant relationships were found between school performance, employment and further education variables and the factor analysis emphasized the importance of 'personality' variables in connection with girls' employment. An attempt was made to assess the value of the various record card items for the prediction of vocational suitability by studying their correlation with the important variable, *employment score*. It was concluded that subject assessments were the most useful and in this connection the 'academic' subjects seemed preferable to the 'practical' ones. Assessments involving groups of subjects were, perhaps, better still and there was evidence to suggest that an overall assessment of performance in a school year (particularly the third or fourth year) might well be the best indicator of all.

I.—INTRODUCTION.

WALKER (1955) reminds us that the cumulative school record has many functions throughout a pupil's school career and at the secondary stage, vocational guidance figures prominently as one of these functions. She also reminds us that the main function of the pupil's cumulative record is to act as a storehouse of detailed information for any one child which no teacher can be expected to remember for a whole class. Obviously, too, staff changes and pupil movement over a period of time may bring about the position that when a pupil reaches school leaving age, the record card is the only point of reference from which we can assess his potential. Certainly, headmasters make use of it for purposes of forecasting. The Youth Employment forms Y 15 are completed from it. It is referred to in the case of any enquiry and also when giving a child a reference. How far, then, are predictions from it valid? Crockett (1945), Vernon and Parry (1949), Walker (1955), and others, tell us to make the school record or extracts therefrom the basis of our assessment when giving vocational advice or predicting the type of job in which a school leaver is most likely to succeed and prompt the question as to which particular aspects of this record are most valuable for the purposes of predicting vocational suitability. The major part of this research endeavours to discover if it is possible to provide a satisfactory answer to this question by assessing quantitatively (as suggested by Rodger, 1951) the various aspects of a pupil's school record card and by calculating the degree of correlation between these aspects and certain employment and further education variables.

II.—PLAN OF THE INVESTIGATION.

(i) *The Samples.*

The pupils chosen for the investigation comprised the first and second year intakes of a large new co-educational secondary modern school in South-East Lancashire, when it opened in September, 1953. These intakes were chosen to allow for at least four years at school (1953-1957) and a similar period of employment and further education (1957-1961). Statistics were collected for samples of 201 girls and 203 boys up to the end of the academic year 1960-61, and were obtained from each of the following major sources:

(1) The scholars' record cards of the school concerned.

(2) Their employment records from the local youth employment bureau.

(3) Examination results of various examining bodies (in particular, the Union of Lancashire and Cheshire Institutes) for the years 1955-1961, covering courses and examinations held in evening institutes, technical colleges, grammar schools and colleges of further education.

(ii) *Summary of the Analyses.*

The investigation was divided into two parts, a general analysis and a quantitative analysis of the variables. The general analysis was made using chi-square and significant relationships between assessments of school performance, attendance, employment and further education were obtained in the case of both boys and girls although the level of significance for the girls' comparisons differed at times from the comparable boys' values. In addition, it was found that a significantly greater number of boys spent time in further education than girls ($P < \cdot 01$). The quantitative analysis which formed the second and major part of the investigation was carried out as follows. The school record (card) was assessed on a five-point scale with an overlap of three grades for each of the three streams. In addition to the *subjects*, four personal attributes (' personality ' variables) were assessed, *appearance* and *conduct* on a five-point scale, *attendance* and *vitality* on a three-point scale. *Employment* was assessed on the final post in terms of the writer's subjective judgment of level of occupation and produced a nine and eight-point scale for girls and boys, respectively. *Movement in employment* consisted of a strict count of the number of moves throughout the pupil's employment record and *direction of movement* was assessed by scoring the first post as well as the final post and subtracting the former from the latter. At the time of this investigation the statutory leaving age for secondary modern school pupils was the end of the term in which they attained their fifteenth birthday. Thus, pupils who remained at school beyond this period or attended full or part-time courses after leaving school were assessed as having spent time in further education on the basis of hours per week, per annum. It will be appreciated, therefore, that *further education time* is a variable applicable to the whole sample, those not taking any further education at all scoring zero. Additional variables involving *results* in further education examinations were obtained for some of the boys and girls who spent time in further education and sat appropriate examinations.

III.—RESULTS.

(i) *The Main Samples.*

The quantitative analysis produced:

(*a*) A matrix of 18 variables for a sample of 192 girls.

(*b*) A matrix of 15 variables for a sample of 190 boys.

These matrices were subjected to a product-moment intercorrelation and the very high correlations obtained between the various *school subjects* suggested that some of them might conveniently be grouped together before submitting the matrices to factor analysis. The groups of subjects were chosen to facilitate comparison between the girls' and boys' samples. This grouping reduced the number of variables in (*a*) and (*b*) from 18 and 15 to 12 and 10, respectively, and the product-moment inter-correlations are shown in Tables 1 and 2.

TABLE 1

THE PRODUCT MOMENT CORRELATIONS FOR THE GIRLS' SAMPLE (N=192) WITH CERTAIN SUBJECTS GROUPED.

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Employment score.		–38	45	48	47	44	39	48	28	34	31	12
2. Employment moves.			–15	–23	–22	–22	–20	–23	–19	–20	–14	–20
3. Further Education Time				37	36	33	28	36	19	15	16	14
4. English, History, Geography, Scripture					98	96	91	95	23	28	25	24
5. Maths, Science, Hygiene						96	90	95	22	26	24	26
6. Needlework, Dom. Science							91	93	21	24	25	24
7. Art								88	17	23	21	18
8. Music									21	29	23	22
9. Appearance										38	05	10
10. Conduct											07	–14
11. Attendance												21
12. Vitality												

(Decimal points omitted.)

For 5 per cent. significance $r = \cdot 14+$ (approx.).

For 1 per cent. significance $r = \cdot 19+$ (approx.).

TABLE 2

THE PRODUCT MOMENT CORRELATIONS FOR THE BOYS' SAMPLE (N=190) WITH CERTAIN SUBJECTS GROUPED.

Variable	1	2	3	4	5	6	7	8	9	10
1. Employment Score		–31	55	58	59	53	38	28	05	20
2. Employment Moves			–17	–21	–24	–22	–15	–25	–09	–18
3. Further Education Time				43	44	38	29	21	11	05
4 English, History, Geography, Scripture					97	91	49	34	25	30
5. Maths, Science, Prac. Drawing						91	48	35	24	27
6. Art							52	31	17	30
7. Appearance								49	12	26
8. Conduct									14	15
9 Attendance										28
10. Vitality										

(Decimal points omitted.)

For 5 per cent. significance $r = \cdot 14+$ (approx.).

For 1 per cent. significance $r = \cdot 19+$ (approx.).

The correlation between *employment score* and other variables was of particular interest and the results for the girls' and boys' samples are compared against this criterion in Table 3, which also shows that the pooled intercorrelations tended to be higher than those between individual subjects.

TABLE 3

THE CORRELATION BETWEEN EMPLOYMENT SCORE AND OTHER VARIABLES AFTER GROUPING CERTAIN SUBJECTS.

GIRLS			BOYS		
Variable	Emp. Score (Criterion)	Ungrouped Correlation	Variable	Emp. Score (Criterion)	Ungrouped Correlation
English, History, Geography, Scripture.	·480	English ·465 History ·491 Script. ·464 Geog. ·461	Maths., Science, Prac. Drawing.	·595	Maths. ·569 Science ·542 P. Dwg. ·618
Music.	·475		English, History, Geography, Scripture	·579	English ·560 History ·579 Geog. ·580 Script. ·551
Maths., Science, Hygiene.	·473	Maths. ·456 Science ·469 Hygiene ·463	F. Ed. Time.	·546	
F. Ed. Time.	·453		Art.	·530	
Needlework, D. Sci.	·441	Nwk. ·435 D.Sc. ·429	Appearance.	·378	
Art.	·389		Employment moves.	–·311	
Employment moves.	–·383		Conduct.	·277	
Conduct.	·339		Vitality.	·198	1% significance.
Attendance.	·307		Attendance.	·052	
Appearance.	·283	1% significance.			
Vitality.	·117				

The importance of *v* factor ability as regards girls' employment is suggested by the subject order of Table 3, whereas for the boys *k* : *m* and *n* factor subjects take precedence.

A comparison of the girls' and boys' inter-correlation, however, shows considerable similarities. High *r* values for the *school subjects*, a marked correlation between the *school subjects* and *employment*, a definite correlation between *further education time* and(in particular) *employment* and *further education time* and the *school subjects*, lower *r* values (particularly, perhaps, for the girls' sample although not in the case of employment) for the *personal attributes* and negative correlations between *employment moves* and the other variables were a feature of both boys' and girls' tables. These similarities tend to emphasize the differences, the more important of which are discussed below.

(*a*) The girls' '*personality*' *variables* correlated more highly with *employment score* than with the *school subjects* suggesting that for girls personality is more important vocationally than educationally.

(*b*) The boys' '*personality*' *variables* correlated more highly with the *school subjects* than with *employment score*. This is exactly the reverse of (*a*) above, and suggests that for boys personality is more important educationally than vocationally.

(*c*) The pattern of correlations between the *school subjects* and *employment score* was considerably higher for the boys' sample than the girls' and emphasizes the importance of intellectual and cognitive ability in connection with boys' employment (cf. Table 3).

(*d*) In the boys' inter-correlation, the numerical, spatial, mechanical group of subjects (*mathematics, science, practical drawing*) appeared to be the most important whereas for the girls it was the verbal factor group (*English, history, geography, scripture*).

(*e*) The correlation between *employment score* and *further education time* was higher for the boys' sample (·55) than for the girls' (·45), which suggests that time spent in further education is more important from the point of view of boys' employment than girls'.

(*f*) The (negative) correlation between girls' *employment moves* and *employment score* (—·38) was higher than the boys' value (—·31). Since movement in employment is an undesirable trait, this would appear to be another aspect of girls' personality having a greater influence as regards employment than boys' personality.

The tendency for groups of subjects to correlate more highly with employment suggested that it would be worthwhile to correlate a *year score*, i.e., a score based on performance in all the record card subjects for any one school year. Scores on this basis were not available for all pupils in the samples and it was necessary to calculate the correlation for each year separately.

The results are shown in Table 4.

TABLE 4

THE CORRELATION BETWEEN EMPLOYMENT SCORE AND THE YEAR SCORES.

	Girls			Boys	
	eta	r_{bis}	N	*eta*	N
Year 1	·44	·53	126	·51	153
Year 2	·50	·61	149	·50	165
Year 3.....	·52	·63	175	·59	154
Year 4	·54	·67	149	·62	167

The *r* values for boys and girls in Table 4 show a similar trend. Whilst direct comparison between the *r* and *eta* values is not possible, there is evidence to suggest (at any rate in the last two school years) that the correlation between *employment score* and the *year scores* is comparable to and perhaps greater than the correlation between *employment score* and *individual subject scores*. Having regard to the fact that increased correlations were obtained by grouping certain subjects we might be justified in assuming that the highest correlations we could expect from an assessment of school performance would be based on groups of, rather than individual, school subjects.

(ii) *The Further Education Sub-Samples.*

(*a*) *Girls.* The girls' further education sub-sample consisted of the eighty-one girls (40·3 per cent.) from the total sample of 201 whose records showed evidence of their having spent time in further education. The amount of time spent by this sub-sample was significantly related to their *employment group*, but not to their *school performance grade*. A detailed analysis of further education examination results showed that sixty-six of the eighty-one girls took some 250 subjects altogether. The examinations were mainly commercial, especially shorthand and typewriting, and the majority were at the first year or preliminary level. Very little advanced work was done except ' O ' level G.C.E., which was taken, in the main, by girls who spent two years full-time at a secondary grammar school after leaving the modern school at 15.

Two additional further education results variables·

(i) a GRADE, indicating a ' general level ' (fail, pass, credit, or distinction) of performance in further education examinations, irrespective of the number taken, and

(ii) a SCORE, based on the actual number of subjects taken and the level of result obtained,

were added to the twelve variables of Table 1 to produce a matrix of 14 variables for the sixty-six girls in the sub-sample. The inter-correlations shown in Table 5 were much lower in value but generally similar in pattern to those obtained for the main girls' sample. The smaller number of girls in the sub-sample, however, raised the value of r required for significance, so that many of the obtained correlations failed to reach the value required for significance even at the 5 per cent. level.

TABLE 5

THE PRODUCT MOMENT CORRELATIONS FOR THE GIRLS' FURTHER EDUCATION SUB-SAMPLE (N=66).

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Employment Score		–26	36	10	15	42	43	39	26	41	24	26	14	–07
2. Employment Moves			–06	–24	–18	–21	–21	–20	–13	–23	–15	–19	17	–16
3. Further Education Time				–07	48	31	30	25	18	23	27	05	04	05
4. Further Ed. Results Grade					34	14	15	07	04	18	11	13	–05	11
5. Further Ed. Results Score						25	27	25	24	24	15	–12	–02	12
6. English, History, Geography, Scripture							97	94	89	95	–07	10	20	–04
7. Maths., Science, Hygiene								95	88	94	–04	14	20	–01
8 Needlework, Domestic Science									89	91	–04	09	23	–03
9. Art										83	–11	03	19	–03
10. Music											–10	12	23	02
11. Appearance												20	02	20
12 Conduct													22	–24
13. Attendance														–03
14. Vitality														

(Decimal points omitted.)

For 5 per cent. significance $r = \cdot 24+$ (approx.).
For 1 per cent. significance $r = \cdot 32+$ (approx.).

We notice from Table 5, that the largest correlation of the girls' further education sub-sample (apart from the subject inter-correlations) is obtained between *further education results score* and *further education time* (·48). The two further education results variables, however, show a negligible correlation with *employment score*, but of the two the ' score ' seems to be the more satisfactory. The non-correlation of *further education results grade* with *further education time* (—·07) may at first seem strange, particularly in view of the high correlation between *time* and *results score* (·48). It should be remembered, however, that the ' grade ' merely indicates a ' general level ' of examination performance which, for any given pupil, is unlikely to be affected by the number of examinations taken (unless adversely, since a high standard is most difficult to maintain) and hence the time spent in preparing for them. The *score* on the other hand, is directly associated with the number of examinations taken and we would thus expect it to be more closely related to the *time* spent in their preparation.

Although the small number of girls (sixty-six) in the sub-sample has made it less satisfactory than it might have been, the overall conclusion to be drawn from the analysis of girls' further education would seem to be that a willingness to spend time at all in further education is, perhaps, the most important factor from an employment point of view. In the main girls' sample we saw fairly high correlations between *further education time* and both *employment* and *school performance* but the amount of time within the sub-sample (i.e., once it is spent) is not so closely associated with *school performance* although it still bears a significant relationship to *employment*. *Results* in *further education* examinations on the other hand, seem merely to be connected with *time spent* which suggests, at least as far as girls' employment is concerned, that to spend time in further education is more important than to take examinations. This, in the opinion of the writer, is what happens in the case of many girls. They take the commercial courses to obtain the practical skill in subjects like shorthand and typewriting, but, because passing at the lower level is not a pre-requisite for entry to the next year's course, they do not, in fact, bother to sit the examinations.

(*b*) *Boys*. The boys' further education sub-sample consisted of the 111 boys (54·7 per cent.) from the total sample of 203 whose records showed evidence of time spent in further education. The amount of *time spent* by this sub-sample was found to have a relationship of very high significance with their *employment group* ($P < \cdot 001$) and a relationship of high significance with their *school performance grade* ($P < \cdot 01$). A detailed analysis of boys' further education results showed that 102 of the 111 boys took 913 subjects in further education examinations and their performance in these examinations had a relationship of very high significance with their *school performance grade*. An analysis of boys' further education examination results is given in Table 6.

Generally speaking, the boys' further education records indicated that most boys took complete (rather than single subject) courses either at an evening institute or on day release or sometimes both. The analysis of their examination results showed a much wider and more varied pattern of courses than those taken by the girls. The analysis in Table 6, however, is of interest, for in spite of the range and variety of boys' courses, over two-thirds of them are at the preliminary or pre-senior (S1) level. According to this table, less than 5 per cent. of the examinations taken reached National Certificate Standard.

The two *further education results* variables, *a GRADE* and *a SCORE* were added to the 10 variables of Table 2 to produce a matrix of 12 variables relating to ninety-seven boys in the sub-sample. The product-moment inter-correlations for this matrix are shown in Table 7.

TABLE 6

BOYS' FURTHER EDUCATION SUB-SAMPLE: AN ANALYSIS OF FURTHER EDUCATION EXAMINATION RESULTS CLASSIFIED ACCORDING TO ACADEMIC LEVEL IN EACH OF THE SCHOOL PERFORMANCE GRADES.

Further Education Examination Standard		Results analysed by School Performance Grades													
revious Classification	Current Classification	A				B				C				Totals	%
		F	P	C	D	F	P	C	D	F	P	C	D		
. Preliminary Courses	C0, C1	5	47	46	18	15	63	32	3	27	42	5	1	304	33·3
. Various	G1, T1, C2, C3, MSE 1	29	91	33	8	28	55	23	9	8	15	7	2	308	33·7
. S1, G.C.E.	G2, T2, C4	19	59	26	1	12	19	10	2	3	2	—	—	153	16·8
. S2	O1, T3	11	42	18	1	4	11	5	1	1	—	—	1	95	10·4
. S3	O2, T4	15	20	3	1	—	—	—	—	—	—	—	—	39	4·3
. Unclassified	—	—	4	—	1	—	—	2	—	5	2	—	—	14	1·5
		79	263	126	30	59	148	72	15	44	61	12	4	913	100·0

F—Fail. P—Pass. C—Credit. D—Distinction.

TABLE 7

THE PRODUCT MOMENT CORRELATIONS FOR THE BOYS' FURTHER EDUCATION SUB-SAMPLE (N=97).

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Employment Score		–24	46	23	36	48	50	41	30	21	–17	–01
2. Employment Moves			–17	–10	–23	–16	–19	–19	–10	–28	–14	–19
3. Further Education Time ..				–06	49	32	34	32	24	05	14	–09
4. Further Ed. Results Grade					39	33	35	28	17	20	–26	06
5. Further Ed. Results Score						32	41	28	26	25	01	–09
6. English, History, Geography, Scripture							95	88	46	18	06	07
7. Maths, Science, Prac. Draw.								87	47	21	03	02
8. Art									50	18	04	13
9. Appearance.............										49	12	10
10. Conduct											10	07
11. Attendance												10
12. Vitality												

(Decimal points omitted.)

For 5 per cent. significance $r = \cdot 20+$

For 1 per cent. significance $r = \cdot 26+$ (approx.).

The pattern of Table 7 follows the now familiar one of high inter-correlations between the *school subjects*, moderate significant correlations between *employment score* and most of the other variables and also between *further education time* and most of the other variables. Again for the boys, as with the girls, value for value, the sub-sample correlations tend to be lower than those for the main boys' sample. The boys' *r* values in Table 7 comparable to the girls' in Table 5, tend to be higher, and (partly because of the lower value required with N=97) many correlations are sufficiently high to be significant at the 1 per cent. level. Both boys' *further education results* variables correlate moderately with certain other variables although the *results score* again seems to be superior to the *grade*. It should also be noted that in the case of the boys, the correlation between *employment score* and each of the *further education* variables is significant, and that the significant correlations between *employment score* and the *further education results* variables is a striking difference between the boys' and girls' further education sub-samples.

Results GRADE would seem a more meaningful variable in the case of boys' further education although still not as useful (judged from the standpoint of correlation) as the *SCORE*. *Results GRADE* however, showed no correlation with *further education time* and the comments in the girls' discussion above are relevant here also. The highly significant correlation between *boys' further education results score* and *employment* (0·36) is probably explained by the demands made by apprenticeships and the necessity (usually) to pass at one level before being allowed to proceed to the next. There was evidence, however, from the analysis of boys' examination results, that whilst this requirement makes the taking of examinations more imperative in the boys' case, it can also be the cause of unnecessary wastage of time. For example, one boy spent three years taking a junior course (Preparatory National Certificate) before he succeeded in passing all three of the subjects at one and the same examination. Another boy taking a Senior Mechanical Engineering Course failed at his first attempt in Engineering Drawing. He took the course again the following year and failed this time in Mathematics. On taking the course a third time he was successful in obtaining a pass at credit standard in all the subjects of the course. Such efforts merit praise for persistence but the time involved must be discouraging to a busy apprentice, especially if it is a matter of one or two marks only, and particularly so when the necessity to pass at the lower level before going on is not always rigidly applied (cf. Wilkinson, 1954, pp. 79-80).

32·8 per cent. of the girls' sample and 50·2 per cent. of the boys' sample took further education examinations. Sixty-six girls took a total of 250 subjects, an average of 3·8 each, and 102 boys took 913 subjects, an average of almost nine each.

IV.—FACTOR ANALYSIS.

The four inter-correlation matrices in Tables 1, 2, 5 and 7 were subjected to a principal components factor analysis and the first four factors were rotated using the varimax technique.

The rotated factors are shown in Tables 8 and 9.

The factor analysis showed quite clearly (particularly in the case of the main samples) that for the boys the *employment score* and *further education time* variables were dependent almost entirely on cognitive efficiency (school performance). On the other hand, the girls' analysis showed *employment score* and *further education time* to be more closely related to the ' personality ' variables suggesting that vocationally, intellectual and cognitive abilities are more important for boys and personality variables for girls. In the case of both boys and

TABLE 8

MAIN SAMPLES · ROTATED FACTORS.

Variable	Girls				Boys			
	Factor				Factor			
	I	II	III	IV	I	II	III	IV
Employment Score	35	50	31	46	79	−05	31	00
Employment moves	−14	−67	01	10	−18	−38	−57	−03
Further Ed Time	28	19	36	52	63	−13	40	−01
English, History, Geography, Scripture ..	95	17	20	06	88	17	04	26
Maths, Science with Hygiene (girls), P.Dwg. (boys)	96	18	11	06	92	14	03	22
Needlework, Domestic Science ...	96	17	06	06				
Art	93	12	05	04	88	13	−02	24
Music	94	20	10	08				
Appearance	09	85	−24	−03	40	15	−11	73
Conduct	09	65	09	24	07	24	12	93
Attendance	19	−15	94	27	02	81	09	10
Vitality	33	20	−15	−35	08	79	08	26
% Variance	40·45	17·31	10·49	6·43	36·38	15·82	6·27	16·51

(Decimal points omitted).

TABLE 9

FURTHER EDUCATION SUB-SAMPLES : ROTATED FACTORS.

Variable	Girls				Boys			
	Factor				Factor			
	I	II	III	IV	I	II	III	IV
Employment Score	35	49	39	18	41	06	61	06
Employment Moves	−15	−11	−03	−75	01	−07	−31	−67
Further Ed. Time	24	82	−10	−19	−07	06	88	06
F. Ed. Results Grade ..	09	03	−09	73	71	01	−01	−04
F. Ed. Results Score ..	26	56	−43	20	21	14	66	−03
Eng. Hist. Geog. Scrip. ..	97	10	06	08	70	37	40	17
Maths. Science with Hygiene (Girls), P.Dwg. (Boys)........	96	12	08	10	69	37	45	13
N. Wk. and Dom. Science	96	09	07	04				
Art	93	−01	−02	−02	67	39	35	22
Music	95	05	07	13				
Appearance	−20	68	15	23	27	79	15	−01
Conduct	01	14	81	29	06	74	01	04
Attendance	23	13	43	−31	−50	44	13	38
Vitality	−04	19	−52	24	15	02	−29	80
% Variance	35·19	12·88	10·78	10·89	20·72	15·23	19·18	11·11

(Decimal points omitted).

(particularly) girls, the ' undesirable ' trait of ***movement in employment*** opposed the other ' desirable ' variables in particular the ' ***personality*** ' ***variables***.*

It was interesting also to note, in the boys' case, that the variable ***further education results grade*** (Table 9) was strongly associated with ***performance at school*** (Factor I) whereas the ***results score*** was more dependent on time spent (Factor III), suggesting, logically, that a boys' ' general level ' of performance (*Grade*) in further education is dependent on his cognitive efficiency, but his overall further education achievement (*Score*) is related to the time he spends on study.

For the girls, the connection between the ***further education*** variables and ***school performance*** was less marked and the ' general level ' of result (***Results Grade***) was more strongly associated with the ' personality ' variables (Factor IV) but for the girls, as for the boys, the ***results score*** was closely related to the ***time spent*** on study (Factor II). Although, from Factor I the importance of a girls' cognitive efficiency as regards employment and further education should not be overlooked, the pattern of the girls' analysis again emphasized the importance of the personality variables in their case (Factors II, III and IV). Interpretation of the girls' analysis (e.g., the bi-polarity of Factor III) was not quite as straightforward as the boys', due perhaps to the smallness of the girls' sub-sample. It was clear, however, from the sub-samples, as with the main samples, that girls' ***personality*** and boys' ***school performance*** tended to be the more important variables as far as ***employment*** and ***further education*** were concerned. In drawing this conclusion, it is not implied that the important connections between school performance and employment and further education which have been apparent for both girls and boys throughout this investigation should be overlooked. The factor analysis, however, has tended to pin-point a fundamental difference between the girls' and boys' analysis by emphasizing a direct connection between ***girls' personality*** and ***employment*** and showing little connection between ***girls' personality*** and ***performance at school***. For the boys the reverse is true. The ***boys' personality variables*** tend to show a definite connection with their ***school performance*** (Factor IV, Table 8), (Factor II, Table 9) but no direct connection with ***employment***.

V.—DISCUSSION.

The results of this research have indicated the existence of a strong connection between a pupil's performance at school (cognitive efficiency) and his subsequent employment. There is also (particularly in the case of the boys) evidence of a marked connection between each of these variables and further education. Nothing in the investigation, however, has suggested that suitability for any particular job can be predicted from any single aspect of school performance, or indeed, by any means whatsoever. One of the difficulties encountered in quantifying and classifying employment was the arbitrariness of placing individual posts into one of a number of groups in some form of hierarchical scale. It has not seemed unreasonable, however, to group posts into broad categories (cf. Reeb, 1959) and this research used nine such groups in the case of girls' employment and eight in the case of boys'. These groups formed the basis of the important variable, ***employment score***. Any attempt, therefore, to assess the most valuable record card items for the prediction of vocational suitability might sensibly be based on the correlation such items have shown with this important employment variable. Judged from this point of view, the following conclusions are suggested :

---

* *Movement* in employment is an undesirable variable since a high score is indicative of instability. The negative correlation of this variable with all others, in particular ***employment score***, is pleasing and supports the findings of Macpherson (1958).

(1) The correlation between employment and the school subjects suggests that this aspect of a pupil's school record is the one likely to be the most satisfactory indicator of vocational suitability. The high subject intercorrelations do not single out any one particular subject as being more useful than any other, but there is evidence that for the girls, the verbal (*English, history, geography, scripture, music*) subjects and for the boys, the mechanical, spatial, numerical (*mathematics, science, practical drawing*) subjects may possibly be superior to the others. In the case of both boys and girls, the 'academic' subjects appear slightly more valuable than the 'practical' subjects.

(2) The grouping of subjects tends to increase the value of their correlation with *employment score* and suggests that an assessment of general performance in a group of subjects might be more useful than any single subject assessment. In this case a *mathematics, etc., group* for the boys and an *English, etc., group* for the girls might be used.

(3) An overall assessment of performance in any particular school year might well be the most useful indicator of all, and in this connection (as might be expected) the third and fourth years at school are more valuable than the earlier ones.

(4) Of the 'personality' variables, *appearance* and *conduct* have appeared to be the most valuable, the former in the case of boys and the latter in the case of girls. However, the importance of personality in the case of girls' employment and further education has been adequately demonstrated. For the boys, personality has its greatest influence on their school performance but as regards boys' employment and further education, cognitive efficiency is all important.

(5) A willingness to spend *time* in *further education* is also likely to be a good pointer to a pupil's suitability for employment because, in the case of boys, *time spent* and *results obtained* are not only closely related to each other but also to performance at school and employment. In the case of the girls, however, the connection between these variables is less marked but there is evidence that a willingness to spend time in further education is more important in the girls' case than the results obtained from such study.

These conclusions, which are merely a summary of the major findings of the research, not only suggest that the pattern of the Youth Employment Service Confidential School Report (Form Y15) might be amended, but endorse the view of Heginbotham, 1951 (p. 184) that these school reports do not give an all-round picture of each pupil and add weight to the suggestion that the Seven-Point Plan (Rodger, 1951) is much more satisfactory. Finally, possible amendments to the present Confidential School Report are discussed, but it is acknowledged that in some instances, further research would be desirable to assess the validity of some of the suggested changes.

ACKNOWLEDGEMENTS.—This article is based on a thesis submitted for the degree of M.Ed. in the University of Manchester, December, 1964. The writer wishes to express his gratitude to Professor Wiseman, for his advice and guidance throughout the research and to the many officials whose permission to examine and use various records made the investigation possible.

VI.—REFERENCES.

CROCKETT, M. (1945). A new post-primary school record *Occup. Psychol.*, Vol. XIX, No. 2, pp. 61-70.

HEGINBOTHAM, H. (1951). *The Youth Employment Service.* London : Methuen and Co.

MACPHERSON, J. S. (1958). *Eleven-year-olds Grow Up.* London : Univ. of London Press.

REEB, M. (1959). How people see jobs . a multi-dimensional analysis. *Occup. Psychol.*, Vol. 33, p. 1-17.

VERNON, P. E., and PARRY, J. B., (1949). *Personnel Selection in the British Forces.* London · Univ. of London Press.

RODGER, A. (1951). *The Seven-Point Plan* . *N.I.I.P. Paper* 1. London : N.I.I.P.

WALKER, A. S. (1955). *Pupils' School Records.* London : Newnes.

WILKINSON, R. L. (1954). Some problems in the education of apprentices, II—Craft training for apprentices. *Vocational Aspect No.* 12, Vol. VI, pp. 78-80.

(*Manuscript received* 8*th April*, 1965)

THE SOCIAL NEEDS AND SATISFACTIONS OF SOME YOUNG PEOPLE

PART I—AT HOME, IN YOUTH CLUBS AND AT WORK.

BY F. MUSGROVE
(*Bradford Institute of Technology*)

SUMMARY. Questionnaires completed by 367 young people between 14 and 20 in the industrial North, indicated the range and nature of the needs which they thought should be met by home, school, club and work, respectively, and the needs which they thought were, in fact, met. Two-hundred-and-fifty subjects were members of mixed youth clubs (fifty were at work, 200 at school). Their responses on a sentence-completion test were compared with the responses of fifty non-club members who were at work and sixty-seven who were at school.

Boys and girls, club members and non-members, young workers and those still at school, all demanded predominantly 'expressive' functions of their homes: feeling at ease, wanted, loved and secure. Approximately a quarter of their demand referred to 'instrumental' functions: character training, instruction in domestic and other skills, preparation for getting ahead in the world. Their demands were broadly satisfied, although some 10 per cent. of club-members' statements referred to frustrations and restrictions at home. Less than 2 per cent of non-club members' statements about home were negative; and it seems possible that club members may often come from comparatively restrictive homes.

The club was expected to provide mainly 'expressive' satisfactions, and did so. Only 6 per cent. of members' statements referred to frustrations. Both work and school were expected to be almost equally 'expressive' and 'instrumental.' The instrumentality of work was seen mainly in terms of character/independence training. School massively failed to fulfill the expressive functions expected of it—particularly the grammar school. But club members were more frustrated by school than non-club members; and again, it seemed that the club might be an important source of 'expressive' satisfactions denied elsewhere. The predominantly negative self-concepts of some grammar-school sixth-formers, who were independently investigated, were in line with the need-frustrations which grammar school pupils appeared to suffer.

I.—INTRODUCTION.

THE needs of young people between 14 and 20 are met by a range of social institutions among which home, school, work and clubs have an obvious importance. The inquiry described below is an attempt to answer two main questions: What are the needs which young people think should be met by home, school, club and work? and, What are the needs which they think are actually met? In answering these questions it was hoped that some light would be shed on the interconnections between institutions, the extent to which one compensates for the deficiencies of another; and that frustrations might be judged by comparing expectations with satisfactions.

It seemed probable that the home might provide relatively trivial satisfactions and a large measure of frustration. Functionalist analyses of contemporary society have often suggested that major satisfactions provided by the home in the past are now provided by other institutions, notably the school (Wilson, 1962); and writing about 'youth culture' often suggests that it originates in conflict with the values and authority of parents. Work was also expected to be a scene of considerable frustration. The mechanization and routinization of work and youthful impatience with authority might be expected to breed widespread resentment among young workers, particularly those in manual employment.

Schools and youth clubs, on the contrary, might be expected to provide important gratifications. The club is a voluntary association and would be unlikely to exist if expectations were widely frustrated. School for pupils over fifteen years of age is also a voluntary organization ; and in any case schools, it is claimed, are today humane, interesting, ' child-centred,' geared with unprecedented psychological insight to pupils' needs ; and they have an obvious value in a relatively open society in improving a young person's prospects in life.

The relationship between personal needs and the ' press ' of institutions has been systematically investigated in American colleges. Job-satisfaction in America has similarly been explored by comparing individual needs with the satisfactions provided by employment (Schaffer, 1953). On a self-rating questionnaire the subject indicates the importance he attaches to allegedly universal human needs such as ' Dominance,' ' Recognition and Approbation,' ' Dependence ' and ' Independence.' He also indicates the levels of satisfaction in these various areas which he derives from his work. Need-strength can be compared with need-satisfaction.

The congruence or incongruence between personal needs and the practices and provisions of American colleges have-been explored by Stern (1962) and Pace and Stern (1958). Again the point of departure is supposed universal human needs such as 'Achievement,' 'Affiliation,' ' Order,' ' Sex ' and ' Understanding.' The psychological needs of subjects are inferred from their responses on an 'Activities Index ' : they indicate their preferences ' among verbal descriptions of various possible activities.' There are descriptions of orderly behaviour, dominant behaviour, deferential behaviour and so on. On the College Characteristics Index subjects score true or false corresponding descriptions of the college environment. Needs can then be compared with the perceived satisfactions, pressures and demands of college life.

But needs are felt in relation to particular institutions. It is true that an institution which manifestly aims to satisfy a particular need may in fact satisfy others. But the boy who goes to a youth club is unlikely to expect satisfaction of the needs which are met through work, or to feel frustrated if these latter needs are not met by the club. He may have a great need for achievement which the club affords little chance to satisfy ; but he will not feel disgruntled, because he never expected it to do so. The author decided to allow spontaneous statements of need in relation to home, club, school and work respectively. No prior assumptions about needs were made : their nature would be determined only after inspection of responses given in an open-ended questionnaire. Satisfactions would be established from corresponding (spontaneous) statements about actual experience of the institutions in question.

II.—METHODS.

Two-hundred-and-fifty members of six mixed youth clubs in a northern conurbation completed usable questionnaires, sixty-seven secondary school children who were not members of youth clubs, and fifty young workers who were not club members. The needs for which subjects sought satisfaction at home were elicited by the following sentence openings:

" At home you *should* always have plenty of chance to . . . "

" First and foremost home *should* help you to . . . "

" At home you *should* always be able to feel that . . . "

The same three openings were applied to school, work and club respectively. Corresponding cues were then provided to elicit statements of actual satisfaction (or frustration) :

"At home you always *have* plenty of chance to . . . "

"Above all else home *does* help you to . . . "

"At home you *do* feel that . . . "

Thus, six statements were made by each subject about his home, his club, his school if he was still at school, or work if he was at work—a total of eighteen statements. Questionnaires were completed anonymously, but classificatory information was obtained and attached to each questionnaire : age, sex, type of education, age of leaving school or proposed age of leaving, examinations taken and passed, professional qualifications, type of employment (for those at work), and father's occupation.

There is an apparent danger that statements of needs are in reality statements of frustration ; that the boy who writes, "At home you should always have plenty of chance to express yourself " says this precisely because he is given no such opportunity. The same difficulty arises in the ' need-press ' analysis of Pace and Stern and the job-satisfaction inquiries of Schaffer. But the latter did not find a negative correlation between need-scores and satisfaction scores, and significant correlations have not been found between corresponding scores on the College Characteristics Index and the Activities Index (McFee, 1961). In the inquiry reported here there was no significant tendency for statements of need to re-appear in negative form in the second part of the questionnaire, as statements of dissatisfaction.

The type of projective test used in the present inquiry has the advantage that it does not present the subject with a perhaps arbitrary and possibly irrelevant list of ' needs ' and ' satisfactions ' which he is forced to reject or endorse. It has been fruitfully employed in America (e.g., Lindgren, 1954) and England (e.g., Bene, 1957) ; and Symonds (1947) found that a sentence-completion schedule not only supplemented personal data obtained in interviews, but was useful in correcting them. (Thus, a report based on interview may describe a man as ' energetic ' ; responses on a projective test may show that a more accurate description would be ' nervous.') It has the disadvantage that the need categories to which responses are assigned cannot be established in advance of the inquiry.

A coarse two-fold classification of responses was made initially into ' expressive ' and ' instrumental ' categories. The distinction is taken from Talcott Parsons : "Action may be oriented to the achievement of a goal which is an anticipated future state of affairs, the attainment of which is felt to promise gratification." " There is a corresponding type on the adjustive side which may be called *expressive* orientation. Here the primary orientation is not to a goal anticipated in the future, but the organization of a ' flow ' of gratifications (and of course the warding off of threatened deprivations) " (Parsons, 1964, pp. 48-9). " Problems of expressive interaction concern relationships with alters which ego engages in primarily for the immediate direct gratification they provide " (Parsons and Shils, 1962, p.209). By extension we refer not only to expressive and instrumental actions and functions, but to expressive and instrumental needs and satisfactions.

After scrutinizing all the responses six subdivisions which appeared logically distinct were made of the instrumental category, and seven of the expressive category. Two judges working independently were able to assign statements to these thirteen subgroups with virtually complete agreement. The instrumental category (I) was subdivided thus: (1) Intellectual skills, understanding and enlightenment; (2) Physical skills (including competence at games and sports); (3) Manual skills (including competence in domestic tasks); (4) Social skills (including poise and self-assurance in relationship with others); (5) Moral development (including references to "forming a good character," "becoming a good citizen," "learning to be self-reliant and stand on your own feet"); (6) Personal advancement (including passing examinations, obtaining promotion, getting on in life).

The expressive category (E) was subdivided into: (1) Ease/emotional security (feeling at ease, wanted, loved, welcome); (2) Freedom/self-direction (including the freedom to express your views, have your say, "be yourself"); (3) Friendship; (4) Sense of competence (including "having a chance to prove yourself"); (5) Support from adults (including "knowing that you can take your problems to parents/teachers/youth leaders"); (6) Sense of identity with the group ("feeling one of the crowd/a member of the family/as if you belong"); (7) Sense of purposeful activity.

When the subjects who took part in the inquiry stated not what their institutions should, but did, provide, statements could be either positive or negative ("At home you always have a chance to relax," "At home you always feel unwanted"). Statements referring to expressive satisfactions were divided into Expressive: Positive (E+) and Expressive: Negative (E –). The latter referred to (1) restrictions, constraints, humiliation, belittlement, rejection, and (2) boredom and demoralization.

The author first administered the questionnaire in six mixed youth clubs which were selected to represent different social areas within a large industrial conurbation. Two clubs were in well-to-do residential suburbs, two in working class districts, and two in socially mixed, transitional, areas, The clubs had a nominal membership of over 300. The clubs were given a week's notice of the author's visit. Club members co-operated well, and over 90 per cent. of those present completed the questionnaire. Two-hundred-and-sixty-eight questionnaires were filled in; eighteen were incomplete, illegible, or otherwise unusable; 250 were used in the analysis.

There were 135 males and 115 females. The age range was 14 to 20; 129 (51·6 per cent.) were 16 years of age or above. One-hundred-and-sixty-three (65·2 per cent.) came from the homes of professional and white collar workers. Two hundred were still at school, 130 at grammar schools and seventy at modern schools. Sixty-seven of the grammar school pupils were in the sixth form. Twenty-four boys and twenty-six girls were at work. Twenty-nine of the workers had attended grammar schools and were all in non-manual, white-collar employment; twenty-one had attended secondary modern schools: the eight boys were all in manual occupations, but ten of the thirteen girls were in routine non-manual employment, mainly as office workers.

For comparative purposes three classes of fourth-year pupils in three secondary schools in the areas served by the clubs were asked to complete the questionnaire. Sixty-seven of these boys and girls were not members of youth organizations. Fifty-seven (85·1 per cent.) came from the homes of manual workers.

Fifty young workers between the age of 16 and 20 were approached in recreational centres. They were not members of youth clubs or formal youth organizations. They were matched for age, sex and type of occupation with the fifty youth club members who were at work. Twenty-two of the twenty-five girls were in white collar employment, fifteen of the twenty-five boys. All fifty completed the questionnaire.

III.—RESULTS.

Home.

Both the expectations and satisfactions of home are predominantly 'expressive.' Two-hundred-and-fifty club members made 750 statements about the needs which they expected home to satisfy, and 77·2 per cent. of their statements referred to expressive satisfactions: being needed, feeling secure, having a chance to talk over personal problems with parents, and the like. Seven-hundred-and-fifty statements were made about the actual satisfactions of home: 72·3 per cent. referred to expressive satisfactions.

In summary, the need-statements were: Expressive (E) 77·2 per cent. Instrumental (I) 22·8 per cent. The statements of satisfaction were: Expressive positive (E+) 72·3 per cent., Expressive, negative (E –) 9·7 per cent., and Instrumental (I) 18·0 per cent.

Boys of all ages and girls over 16 place more stress on the instrumental functions of home than do younger girls. 29·8 per cent. of older girls' demands referred to instrumental needs, only 15·3 per cent., of the younger girls' demands (CR 3·9, $P<0{\cdot}001$). The younger girls also find more dissatisfaction at home: 15·3 per cent of their statements refer to restrictions and other grievances, only 6·4 per cent of the older girls' statements (CR 8·3, $P<0{\cdot}001$).

Although there is a broad correspondence between needs and satisfactions, there are some major shifts within sub-categories. While the greatest demand is for emotional security (approximately a third of all need-statements), this is closely followed by the demand for freedom and self-direction (approximately a quarter of all statements). The first demand is met in full, the second falls substantially short of satisfaction. Twenty-two point five per cent. of the boys' demands and 23·9 per cent of the girls' referred to freedom and self-direction at home; only 13·3 per cent. of the boys' stated satisfactions mentioned freedom and self-direction, and only 11·3 per cent of the girls' statements.

Nevertheless, home emerges on this sentence-completion schedule as a pre-eminently satisfying social institution for young people. This is in line with other inquiries into home life today. Inquiries among representative samples of English adolescents (Hancock and Wakeford, 1965), as well as comparative studies sponsored by Unesco (Berge, 1964), indicate close bonds between young people and their parents and widespread appreciation of parents and reliance on them as guides and counsellors in times of trouble.

In the present inquiry tributes to parents were frequent at all social levels. Subjects often stated a desire to talk over problems with their parents, and as often stated that they were in fact able to do so. A 15-year-old daughter of a plumber wrote: "At home you always feel you can tell your troubles to your parents and not be laughed at"; a 15-year-old son of a labourer: "At home you always have a chance to tell some-one your troubles"; a 16-year-old apprentice hair-dresser, daughter of a textile buyer: "At home you always have plenty of chance to talk sensibly to your parents"; a 17-year-old postman's daughter: "At home you always feel that your opinions matter"; the 15-year-old son of a tailor: "At home you always have plenty of chance to talk over matters with your parents" ; the 15-year-old daughter of a fireman : "At home you always have plenty of chance to air your views."

Appreciation of home was expressed in more general, and often enthusiastic, terms. The 18-year-old daughter of a clerical worker wrote : "At home you always feel confident that you are needed by some-one" ; the 16-year-old son of a shop manager : "At home you always feel joyful" ; the 14-year-old son of a bus driver : "At home you always feel wanted and grateful to your mother." Home offers relaxation, emotional security, and freedom to "be yourself." The 14-year-old daughter of a shopkeeper wrote : "At home you always feel relaxed, unafraid, easygoing, and enjoying every minute of it." (The same girl says of her grammar school : "You always feel petrified of punishment and nervous of exams.") The 17-year-old son of a company director claimed that : "At home you always feel happily unconscious of life." ; the 16-year-old daughter of an engineer : "Above all home helps you to have roots in a crazy world." An 18-year-old typist, daughter of a civil servant, said that : "At home you always have plenty of chance to relax after work and you do not feel that you have to put on an act."

At home you can be natural ; individuality is recognized ; a person is not an anonymous member of a larger social category. "At home," says the 14-year-old son of a manufacturer, "you should be able to feel an individual in your own right." Home should be, and is a place where your opinions count. The 16-year-old son of a businessman was of the opinion that "At home you should have plenty of chance to take part in things and have a say in things—like interior decoration (the colour of the wallpaper), to take one of many examples." Such demands appear to be met. "At home," says the 17-year-old son of a salesman, "you do feel that it really is a home, and not just a house you happen to live in."

The instrumental purposes of home are overshadowed by the expressive ; but more than 10 per cent. of the expectations of home refer to moral training and character development, "learning to stand on your own feet," "becoming a responsible citizen," "having a more sympathetic outlook toward others." Table 1 shows the level of demand and satisfaction in this regard.

TABLE 1

STATEMENTS ABOUT MORAL DEVELOPMENT
(Boys and Girls).

	Demand	Satisfaction
	%	%
School (N 600) ..	9·8	6·0
Club (N . 750)	4·0	4·4
Home (N · 750)	12·1	7·8

Social skills, manual skills and intellectual enlightenment are also sought at home, but to a smaller extent. About three per cent. of the subjects looked for some form of intellectual understanding. "Above all," said a 17-year-old girl, "home helps you to understand about adult things." A 17-year-old boy thought that "First and foremost home should help you to secure knowledge of things not taught at school, such as sex." A general social competence is sometimes mentioned: "Above all home helps you to manage your money and look after yourself." (18-year-old typist). "Above all home helps you to learn how to live with people and how to be well-mannered" (18-year-old daughter of a clerical worker). The 14-year-old son of a bus driver claimed that "Above all home helps you to learn a lot about gardening." For girls the instrumentality of home lies mainly in learning domestic skills: "home helps you to learn about married life," and "home helps you to learn how to run a home."

9·7 per cent. of statements about the actual experience of home expressed dissatisfaction or criticism (8·4 per cent. of the boys' statements, 11·0 per cent. of the girls'). Some references were made to being misunderstood (a 15-year-old girl, daughter of a Y.M.C.A. secretary, said: "At home you feel that you're always in the wrong," the 15-year-old son of a bingo hall manager said that at home "you always feel that everyone is against you"). In the main grievances referred to restrictions and an overload of chores. "At home you always feel that you must do a few jobs and bring home an unopened wage packet" said an 18-year-old laboratory assistant. An 18-year-old tripe dresser always has the urge, when at home, "to be off all the time on my motor-bike." A 16-year-old trainee draughtsman "always feels cooped up" at home; a 15-year-old engineer's daughter "always feels in the way"; the 16-year-old daughter of a civil servant "always feels helpless." A 15-year-old girl, daughter of an insurance agent, always "feels like an unpaid char." Some—by no means entirely from working class homes—complain of the lack of privacy at home. And the mass media are intrusive. Although some appreciate home as a place where television is constantly available (a 14-year-old boy appreciates the chance to "loll around all the time with a cup of tea watching television"), others are like the 17-year-old son of a tailor who feels, at home, that he is "televisionized, radioised, bookised, and newspaperised."

Home is an enclosed world, isolated from wider social contacts, and largely accepted as such. The girls make some demand that they should be able to meet their friends at home (3·8 per cent. of responses), but only 2·1 per cent. mention home as a place where they do in fact meet their friends and non-family persons. Only one per cent. of the boys mention home as a place where friends should or can be met. It is pre-eminently the club that is expected to meet the need for non-family contacts; and the need appears to be met in full.

The Youth Club.

Like home, the youth club is seen primarily in 'expressive' terms. It is not valued for any training it may provide in intellectual or even physical skills; in so far as it is instrumental, it is required to provide social training.

The needs which subjects hoped the club would satisfy were: 82·9 per cent. expressive, 17·1 per cent. instrumental. The actual satisfactions were: E+ 78·3 per cent., E − 5·9 per cent, and I 15·8 per cent.

When the club was mentioned as a place to meet friends and be friendly with people, the statement was placed in the expressive category. ("Expressive action is not oriented to the attainment of a goal outside the immediate action situation and process itself in the same sense as is instrumental action" (Parsons, 1964).) When there was clearly a notion of learning to get on with people, developing one's social skills, the statement was placed in the instrumental class. In such statements the club was seen as a training situation, helping you to overcome shyness, gain social confidence and poise, learn how to mix and get on with people.

Above all club members demand from their clubs a sense of ease and emotional security. A quarter of their statements refer to feeling relaxed, at ease, wanted, welcome. A comparable number of statements refer to being friendly, meeting and making friends. Other statements refer specifically to a sense of group solidarity, feeling ' part of it,' being ' one of the crowd.'

TABLE 2

FRIENDSHIP AND THE SENSE OF BELONGING

(Boys and Girls).

	DEMAND				SATISFACTION			
	School (N 600)	Club (750)	Home (750)	Work (150)	School (N . 600)	Club (750)	Home (750)	Work (150)
	%	%	%	%	%	%	%	%
Group solidarity	2·0	7·6	7·7	10·0	2·3	5·3	7·1	9·0
Ease and security ...	8·0	24·9	34·3	6·9	1·5	28·4	33·8	8·3
Friendship	4·7	23·8	2·3	8·6	4·5	21·7	1·6	9·3
Totals	14·7	56·3	44·3	25·5	8·3	55·4	42·5	26·3

The demand for friendship often makes explicit mention of the opposite sex. Demands for specifically sexual contact or experience were rare, made by half a dozen boys. A 16-year-old grammar school boy thought that "At the club you should have plenty of chance for sex" ; another said that "At the club you always feel sexy" ; and a third 16-year-old boy said that "At the club you always feel in a good and sexy mood." Other-sex references were usually in less specifically sexual terms. A 14-year-old boy always feels happy at the club " because of Valerie, Susan and Joan." Sex antagonism as well as attraction was mentioned by three boys, thus : "At the club you always feel opposed to the opposite sex " (a 16-year-old grammar school boy).

The club is expected to provide a place of refuge from the stress of life outside. Its main purpose is to offer compensation for the restrictions, anxieties and humiliations promoted by other institutions. It is an escape from the constraints of school and the surveillance of parents. "At the club," says the 14-year-old son of a solicitor, " you should be able to feel relaxed and forget the terrible turmoil and distress of the world." "Above all the club helps you to get out of the school routine " (14-year-old grammar school girl).

Club membership legitimizes absence from home. "Above all the club helps you to go out without my dad arguing " (14-year-old daughter of a lorry driver). It is an escape from parents : "Above all the club helps you to do something constructive without criticism from parents : it is completely separate from parents " (17-year-old daughter of a senior civil servant). "At the club you always feel that there is no-one there to snap at you " (14-year-old daughter of a lorry driver). "At the club you should always be able to feel relaxed, without grown-ups nagging on to do this and not to do that" (15-year-old labourer's son). "Above all the club helps you to get away from the washing up " (14-year-old architect's daughter).

At the club you should be able to express your views, have your say, put forward a point of view. Thirteen per cent. of all demands were of this character ; and 11 per cent. of stated satisfactions were in these terms. Still greater demands of this nature are made of school and home, but satisfaction falls much further short of demand.

The club is expected to a greater extent than school or home to provide training in social skills. " First and foremost the club should help you not to be afraid of meeting people " (14-year-old modern school girl, daughter of a shopkeeper). " The club should help you to talk to people more easily than you would otherwise " (15-year-old son of a chemist). " First and foremost the club should help you to socialize " (17-year-old grammar school boy).

TABLE 3
STATEMENTS ABOUT SOCIAL SKILLS.

	DEMAND	SATISFACTION
School	%	%
Boys (N · 336)	2·9	1·8
Girls (N : 264)	7·2	3·4
Club :		
Boys (N : 405)	8·1	5·2
Girls (N : 345)	10·7	7·8
Home :		
Boys (N 405)	2·7	1·0
Girls (N 345)	2·0	1·5

Although the club is appreciated for the informal training in sociability which it affords, complaints are made of undue pressure from the leaders to ' mix in ' and conform. The 16-year-old daughter of an engineer says that at the club she feels " organised, frustrated, got-at, expected to conform, keep quiet, take part in ridiculous discussions, able to hike." The 15-year-old daughter of a fishmonger feels that " you're always expected to go along with everyone else " ; the 15-year-old son of a bricklayer that " you're always expected to do what the others are doing."

Club leaders hope that the club will be far more instrumental than members want it to be. The questionnaire for members was modified for leaders. Twenty-seven leaders were asked to say what young people should have a chance to do at the club, in what ways the club should help them, and what, above all, they should feel when they were at the club. They were then asked what they thought, from their experience, young people did in fact look for in the club. They placed considerable stress on instrumental, mainly moral purposes ; but when they considered reality, they recognised that ' expressive ' purposes were predominant. (No club leader made an adverse comment on reasons for attending the club.)

TABLE 4
FUNCTION OF THE YOUTH CLUB
Statements by Leaders and Members

	SHOULD		IS		
	E	I	E+	E−	I
	%	%	%	%	%
Leaders (N . 81)	53·1	46·9	74·0	nil	26·0
Members (N . 750)	82·9	17·1	78·3	5·9	15·8

The 'instrumentality' that leaders hope for in the club is chiefly in the form of character training: teaching members to accept responsibility, to become useful citizens, to find a true sense of values, to achieve 'maturity.' They recognize that they come in the main to feel welcome, at ease, unconstrained and relaxed.

Of course some members valued the club in precisely the instrumental terms that leaders would approve. "Above all," said a 17-year-old grammar school boy, "the club helps you to develop your administrative powers." "Above all, the club helps you to learn self-discipline and human relations" (another 17-year-old grammar school boy). But the predominant demand was for current feelings rather than for future competence and personal development.

Statements expressing dissatisfaction with the club were few: 5·9 per cent. of all statements made. Since the club is a voluntary association, this low proportion is not, perhaps, surprising. Dissatisfaction was almost wholly directed towards aimlessness and lack of clearly defined purpose: "At the club you feel that you never do anything out of the ordinary and it is often boring."

Work.

In the sentence-completion schedule completed by the fifty youth club members who were in full-time employment, work appeared to be generally satisfying, and in unexpected directions. The number of subjects in this group is, of course, small; and sub-groups (white-collar female workers, ex-grammar school boys, etc.) are too small to make reliable comparisons possible.

As one would expect, work was seen in instrumental terms to a far greater extent than home or club. But work was an instrument for other than crudely material ends: in fact, it is referred to largely as a learning situation (in a broad sense) by both manual and non-manual employees.

Forty-eight per cent of subjects' statements referred to expressive needs, 52 per cent. to instrumental needs. Stated satisfactions were: E+ 45·3 per cent., E − 7·3 per cent., and I 47·4 per cent.

Boys and girls are alike in placing great stress on the educative value of work: understanding not only of techniques and processes, but of people and wider social problems should be promoted. Boys differ from girls in placing considerable stress on work as a means to "getting on in life." Girls look for friendship at work; but boys also place a high value on a sense of solidarity with the work group or organization.

"At work you always have plenty of chance to learn something new and to ask about things without being considered a nuisance," says a 17-year-old trainee nurse. But it is not only trainees who appreciate work as education; indeed, for a 16-year-old apprentice electrician, work seems to mean, above all, a place "where you are expected to put the kettle on at 9-55 sharp."

It is the wider education in human affairs and understanding that receives most frequent mention: "Above all work helps you to increase your knowledge of society" (16-year-old audit clerk); "At work you always have plenty of chance to see how people react under the same conditions" (18-year-old male clerk); "At work you always have plenty of chance to meet people and learn about different people" (15-year-old grocery assistant); "Above all work helps you to be able to listen to older people's ideas and views" (18-year-old typist). And work promotes skill in human relationships: "Above all work helps you to get on with other people" (18-year-old printing works operative).

Work is expected to develop ' character.' No less than 21 per cent. of the demand is for moral training, in a broad sense. (On the other hand, only 11 per cent. of subjects' statements about reality refer to satisfactions in this regard.) Work is expected to help you to become responsible, to learn to shoulder responsibility, to help you to " grow up " to learn to " stand on your own feet." And in large measure it is seen as promoting a sense of responsibility, independence and self-reliance. Work " helps you to prepare for responsibility later in life," says a 17-year-old office worker ; " Above all work helps you to become a more mature person " (17-year-old fitter) ; "Above all work helps you to rely on yourself " (17-year-old male shop assistant) ; "Above all work helps you to have a life of your own, giving more independence than school did " (ex-grammar school clerical worker).

15 per cent. of the boys' statements demanded a sense of corporate membership at work ; 11 per cent. of their statements of satisfaction indicated that they achieved this sense of identification. (Only 5 per cent. of the girls' demands and satisfactions referred to this sense of belonging to an organization or group.) "At work you always feel part of the firm," said a 19-year-old articled clerk ; "At work you always feel you are an essential part of the mechanism " (16-year-old bank clerk).

7·3 per cent. of the statements describing actual experience of work referred to frustrations and dissatisfactions. In no case was there any complaint about pay (except an oblique reference by a Post Office maintenance engineer : "At work you always feel that it is a rich man's world ").

A minority of statements referred to a sense of wasting time or being under undue pressure (" always feeling rushed off your feet "). A few statements referred to a sense of humiliation : "At work you feel that some adults do not treat young people with any kind of respect " (18-year-old ex-grammar school girl employed as a clerk). "At work you always feel you're being told what to do " (18-year-old operative in a printing works). But the broad picture which emerges is one of appreciation of work as an environment which fosters growth towards a self-confident and self-respecting maturity.

IV.—REFERENCES.

Bené, Eva (1957). The use of a projective technique, illustrated by a study of the differences in attitudes between pupils of grammar schools and secondary modern schools. *Brit. J. Educ. Psychol.*, 27, 89-100.

Berge, A. (1964). Young people in the Orient and Occident. *International J. of Adult and Youth Educ.*, 16.

Hancock, Alan, and Wakeford, John (14th Jan, 1965). The young technicians. *New Society*, 120, 13-14.

Lindgren, H. C. (1954). The use of a sentence completion test in measuring attitudinal change among college freshmen. *J. Soc. Psychol.*, 40, 79-92.

McFee, Anne (1961). The relationship of students' needs to their perceptions of a college environment. *J. Educ. Psychol.*, 52, 25-29.

Pace, C. Robert, and Stern, George, G. (1958). An approach to the measurement of college environments. *J. Educ. Psychol.*, 49, 269-277.

Parsons, Talcott (1964). *The Social System.* London : Routledge.

Parsons, Talcott and Shills, Edward (1962). *Toward a General Theory of Action.* New York : Harper Torchbooks.

Schaffer, Robert H. (1953). Job satisfaction as related to need satisfaction at work. *Psychol. Monogr.*, 67,14.

Stern, George G. (1962). Environments for learning. In Sanford Nevitt : *The American College.* New York : John Wiley.

Symonds, P. M. (1947). The sentence completion test as a projective technique. *J. of Abnorm. and Soc. Psychol.*, 42, 320-329.

Wilson, B. R. (1962). The teacher's role—a sociological analysis. *Brit. J. Sociol.*, 13, 15-32.

(*Manuscript received 29th April*, 1965.

A COMPARISON OF THE PERFORMANCE OF SUMMER AND AUTUMN-BORN CHILDREN AT ELEVEN AND SIXTEEN

BY H. G. ARMSTRONG
(Educational Psychologist, West Riding)

SUMMARY. The performance of a complete age group of some 24,000 children on standardised tests at the age of 11 was examined to determine whether they discriminated in favour of Autumn, as compared with, Summer-born children. No significant difference was found. A similar comparison was made at the Ordinary Level of the General Certificate of Education and again no difference in performance was evident. This obtained no matter what form of organisation prevailed in the school. A further comparison was made between children in urban and rural areas at 11 under test-free conditions. This revealed that Summer-born children in large urban schools were at a disadvantage in comparison with those in small rural schools.

I.—INTRODUCTION.

A CONSIDERABLE body of evidence is now available in relation to the effects of (*a*) season of birth on intellectual performance and (*b*) of school organisation on children born in the Summer and Autumn. It is with the latter question that this investigation is concerned.

The earlier evidence on seasonal changes has been summarized by Pidgeon and Dodds (1961). Johns (1962) draws attention to the number of Summer-born children retarded in reading in the junior school, as compared with the Autumn-born, and the high proportion to be found in lower forms. This is confirmed by Jinks (1964). " As a result of whatever form or system of grading the junior school uses to stream its intake, the children with the most schooling behind them perform best and, as a result, are placed in the 'A' class." There is also, according to Jinks, a marked effect on 11+ allocation, in spite of the allowance made for age in standardised tests. " Of the 317 children entering grammar schools, 120 were born in September to December, 103 in January to April and 92 in May to August." These, however, relate to a comparatively small age group and, since school reports and grading assessments are also taken into account in the area from which Jinks's figures are drawn, it is not known to what extent these bias the final sleection in favour of the older children.

Jackson (1964) has pointed out that as schools increase in size the discrepancy between Summer and Winter-born children increases. In four-stream schools almost twice as many Summer-born as Winter-born children are to be found in the ' D ' stream. Williams (1964), from a study of 265 children in eight schools for the educationally subnormal, demonstrates that the number of Autumn births falls below, and the number of Summer births exceeds, expectation. Of the two factors affecting the situation, namely term of entry and age-group position, Williams points out that the Summer-born children more often retain their position as youngest in the age group and that teaching levels may not take this sufficiently into account. Although he could not separate the effects of ' term of entry ' and ' age-group position,' his data suggested that the latter may be the more powerful factor. Evidence from Sweden, where the complete age group starts schools at the beginning of the school year, points to the operation of the 'age-group position' effect. Johansson (1965) has demonstrated that both assessments and tests of school readiness give low but significant correlations with age at start of school.

II.—PERFORMANCE AT ELEVEN-PLUS.

The present study is concerned mainly with children selected for grammar schools and compares the performance of Summer and Autumn-born children at 11 and again at G.C.E. level.

Table 1 below gives the quarterly proportions selected at 11 from a complete age-group :

TABLE 1

PERCENTAGE OF QUARTERLY AGE GROUP SELECTED FOR GRAMMAR SCHOOLS.

	Monthly Quarter	No. in Age Group	No. Selected	Quarterly Proportions
1	June—August ..	6903	1543	22·3
2	March—May ..	6192	1451	23·4
3	Dec.—Feb. ..	5638	1277	22·6
4	Sept.—Nov. ...	5413	1192	22·0
	Total	24146	5463	22·6

These results, based on Moray House Tests of English, Arithmetic and Verbal Reasoning, show that the proportions selected vary very slightly from one quarter to another, the biggest difference, 1·4 per cent., arising between the Spring and Autumn-born children.

Table 2 indicates whether there is any significant difference as a result of this variation.

TABLE 2

OBSERVED AND EXPECTED DISTRIBUTIONS.

	1	2	3	4
Observed............. .	1543	1451	1277	1192
Expected................	1562	1404	1273	1224

The distribution is not significant at the 5 per cent. level.

$\chi^2 = 2{\cdot}65$. For 3 d.f. $0{\cdot}30 < p < 0{\cdot}50$.

The contribution of the Summer and Autumn born to chi-squared is less than ·50 per cent.

III.—PERFORMANCE OF SUMMER AND AUTUMN-BORN CHILDREN IN G.C.E.

Table 3 compares the performance of those at opposite ends of the age range from approximately sixty grammar schools in the Ordinary Level examination of the G.C.E.

TABLE 3

G.C.E. RESULTS.

	No. entered	Average No. of subjects passed	S.D
June—August	1028	4·25	2·53
Sept.Nov. .. .	1059	4·41	2·52

Table 3 shows that there is no significant difference in performance between those who fall in the youngest and oldest quarters of the age group. $\frac{D}{\sigma_D} = 1{\cdot}45$

Similar details are given in Table 4 for thirty-two grammar schools in which the entry is streamed on admission.

TABLE 4

RESULTS IN STREAMED GRAMMAR SCHOOLS.

	No. entered	Average No of subjects passed	S D.
June—August	549	4·15	2·42
Sept.—Nov.	539	4·34	2·44

Again, there is no significant difference. $\frac{D}{\sigma_D} = 1{\cdot}36$. It appears that, no matter what form of organisation prevails, there is no difference in performance between Summer and Autumn-born children at G.C.E. level. For schools unstreamed in the first year (N=10), $\frac{D}{\sigma_D} = 1{\cdot}57$ and for schools where 'setting' obtains throughout (N= 5), $\frac{D}{\sigma_D} = {\cdot}40$.

IV.—SCHOOL ORGANISATION.

Research evidence, nevertheless, clearly shows that the younger children of an age group are likely to be at a disadvantage when primary schools are streamed, that the incidence of learning failure is greater amongst them and that they are to be found in E.S.N., schools in greater numbers than expected. It is probable that any modification of school organisation which breaks down group instruction is of advantage to the younger children. Tables 5 and 7 below, which contrast the performance at 11 under test-free conditions of children in urban and rural areas, lend support to this contention :

TABLE 5

URBAN AREAS.

Monthly quarter	No. in age group	No. selected	Quarterly Proportions
1	480	131	27·3
2	547	170	31·1
3	438	143	32·6
4	466	160	34·3
Total	1931	604	31·3

These results are derived from eighteen urban schools in two consecutive years.

TABLE 6

OBSERVED AND EXPECTED FREQUENCIES.

	1	2	3	4
Observed	131	170	143	160
Expected	153	169	137	145

Here the Summer and Autumn-born account for 95 per cent. of the total value of chi-square.

$\chi^2 = 4{\cdot}98$. For 3 d.f. $0{\cdot}10 < p < 0{\cdot}20$.

TABLE 7

RURAL AREAS.

Monthly quarter	No. in age group	No.s elected	Quarterly Proportions
1	396	85	21·5
2	449	100	22·3
3	391	93	23·8
4	344	70	20·4
Total	1580	348	22·0

The above results are drawn from eighty-eight rural schools during three consecutive years of a test-free scheme.

TABLE 8

OBSERVED AND EXPECTED FREQUENCIES.

	1	2	3	4
Observed	85	100	93	70
Expected	87	99	86	76

The distribution is clearly in accordance with expectation.

$\chi^2 = 1{\cdot}10$. For 3 d.f. $0{\cdot}70 < p < 0{\cdot}80$.

The main point at issue is the effect of size of school. In the one case the average size of age group exceeds fifty, in the other it barely reaches six. Had both areas taken the same standardised tests, the differences revealed by size of school in a test-free scheme would have been masked.

V.—CONCLUSIONS.

(1) The performance of a large age group at 11 does not support the view that standardised tests discriminate against Summer-born children.

(2) Whether a child spends one or two terms more or less in the infants' school makes no difference to his ultimate achievement.

(3) Whether a grammar school is streamed, unstreamed or organised on the basis of ' setting,' there is no appreciable difference in attainment between the Summer and Autumn-born as measured by G.C.E. results.

(4) A comparison of urban and rural areas under test-free conditions suggests that by itself unstreaming of urban primary schools will not remove any handicap borne by the youngest children unless accompanied by individual methods of teaching.

VI.—REFERENCES.

DAVIES, A. M. (1964). Season of birth, intelligence and personality measures. *Brit. J. Psychol*, 55, 475-6

JACKSON, B. (1964). *Streaming : An Educational System in Miniature.* London : Routledge and Kegan Paul.

JINKS, P. C. (1964) Investigation into effects of date of birth on subsequent school performance *Educ. Research*, Vol 6, No. 3.

JOHNS, E. (1962). The age-factor in reading retardation. *Researches and Studies*, Vol. XXIV, Univ. of Leeds Inst. of Education.

ORME, J. E. (1963). Intelligence, season of birth and climatic temperature. *Brit. J. Psychol.*, 54, 273-6.

PIDGEON, D. A., and DODDS, E. M. (1961). Length of schooling and its effects on performance in the junior school. *Educ Research*, Vo. III, No. 3.

WILLIAMS, P (1964) Date of birth, backwardness and educational organisation. *Brit. J. Educ Psychol.*, 34, 247-55

(*Manuscript received 20th September*, 1965)

A STUDY OF DIFFERENCES IN THE JUDGMENTS OF ADOLESCENT PUPILS

BY E. A. PEEL

(University of Birmingham)

SUMMARY. A cross-sectional method was devised to investigate the differences in the judgments and reasoning of adolescent pupils when presented with connected passages calling for comprehension. Two passages were written to bring out certain features of the pupils' thinking such as tautological and irrelevant thinking, circumstantial judgments and finally, comprehensive judgments involving the imagination and evocation of possible explanations. Three large groups of pupils were tested, representing a wide range of intellectual capacity and achievement. The pupils' answers to the questions set on the passages were capable of analysis in terms of categories ranging from pre-logical answers to fully comprehensive answers evoking imagined possibilities. Interesting formal changes were also demonstrated. In general, pupils up to the age of 13½ years judged circumstantially and only by 14+ years did they show a firm tendency to make comprehensive judgments involving the production of possible explanations.

I.—THE PROBLEM.

ALTHOUGH much attention is being given to the study of the growth of children's thinking, most of it is concerned with their reasoning in arithmetic, logic, mathematics and science and most of the situations involve judgments about non-verbal material—pictures, perceptual forms, concrete objects and science experiments. Some work has also been done by using literal codes (Peel, 1960, p. 107). However, the changes in verbal comprehension during adolescence have received little attention, although a useful start was made on the study of children's reasoning from short passages of prose in terms of Piaget's theory (Lodwick, 1958; Peel, 1959), the technique being extended later to the study of the thinking of secondary school pupils and students at a college of further education (Case and Collinson, 1962).

A study of the changes in verbal comprehension and thinking is important since nearly all school material in history, English and allied subjects consists of connected prose. To a considerable extent also, progress in geography and the sciences depends on the understanding of ideas and reasoning expressed in the same form. Furthermore, the changes that take place in adolescent understanding and judgment have not been fully investigated, though they are often taken for granted in educational reports (Crowther, 1960; Newsom, 1963).

During the last few years attempts have been made to probe more thoroughly and perceptively into the changes in thinking during adolescence (Peel 1965). This paper describes the basic material used and presents the results obtained from it. The type of material used and the method of analysis devised have also been successfully extended more recently to the study of thinking in geography, history, current affairs and general comprehension of connected material (Peel, 1965; Rhys, 1966; Davies, 1965; Stones, 1965.) Pupils' religious understanding has been similarly investigated (Goldman, 1965).

II.—THE MATERIAL.

The main task was to construct material and questions which enabled one to concentrate upon the quality of children's judgments and the changes taking place in them over the critical years between 9 and 15. It was hoped not only to demonstrate clear-cut differences between the thinking of the 10-year-old and the 15-year-old pupil, but also to reveal what finer changes may take place between the extremes.

The material consists of two short anecdotes, each followed by an open question about some feature of the passage and a second question asking for the reasons for the answer given to the first question. The two passages consist of *logically* similar material.

The passages and questions are as follows :

(i) Jane is a very clever 15-year-old girl who is preparing for her final examinations· One evening, as Jane was doing her homework, her mother asked her to look after her younger brother, Teddy, while she went out. Teddy wandered from the living room into the kitchen, got hold of a jar of jam, ate a lot of it, covered his clothes with it, and spilled it on to some clean washing in a basket so that it was ruined.

Now answer the following questions :

(1) Was Jane a careless person ? ....

(2) Why do you think so ? ....

(ii) Only brave pilots are allowed to fly over high mountains. This summer a fighter pilot, flying over the Alps, collided with an aerial cable-way, and cut a main cable, causing some cars to fall to the glacier below. Several people were killed and many others had to spend the night suspended above the glacier.

(1) Was the pilot a careful airman ? ....

(2) Why do you think so ? ....

Each passage contains a leading statement which is irrevelant to the judgment required in answering the question put to the pupil, i.e., *Jane is a very clever 15-year-old girl . . .*, *Only brave pilots are allowed . . .* Then follows a short section in which a happening takes place which is related to the question, but which is not by itself adequate or sufficient to form the basis of a judgment. In the one case the happening is *what befalls Teddy* and in the other *the collision with the cable*. A limitation is also imposed by other circumstances stated or implied. The limitations are first those connected with the stated fact that Jane is busy at her homework, and secondly those implied by the hazards of flying over the Alps.

Thus, responses could be evoked at least at three points on a scale of maturity of judgment :

(*a*) Where irrelevancy, tautology and inconsistency may dominate.

(*b*) Where the content or circumstances (Teddy and the jam, cutting the cable) solely may decide.

(*c*) Where extenuating possibilities are invoked, and where all elements in the situations are explored in support of the judgment made.

III.—SCHEME FOR ANALYSING THE REPLIES.

The basic scheme for classifying the responses is that set out under (*a*), (*b*) and (*c*) above, with detailed elaborations to be described later, but, since the replies to each test refer to its specific content, the scheme of classification will be set out separately for each anecdote.

Jane.

Q : Was Jane a careless person ?

Five categories of answers were selected which were as follows beginning with the category representing the most comprehensive judgment.

(Category cii) No, Yes, Maybe, not sure, taking account of extenuating possibilities (importance of examinations, mobility of younger brothers, impossibility of the commission).
e.g., No. She had a lot of work to do for her examination.
Younger brothers are hard to watch.

(Category ci) Yes, although she was doing her homework, she should have watched her brother.

(Category bii) Yes, she should have watched Teddy. If she wasn't careless, she would have watched Teddy ; *or* If she had watched Teddy he would not have wandered.

(Category bi) Yes, because she let Teddy wander.

(Category a) No, with irrelevance (she's clever) denial of terms (it's Teddy who was so careless).

The Pilot.

Q : Was the pilot a careful airman ?

Four grades were chosen (there being no instances of Category ci answers) :

(Category cii) Yes, No. Maybe, taking account of extenuating possibilities, vision, weather, state of the plane.

(Category bii) No, because if he was careful, he would not have cut the cable.

(Category bi) No, because he hit the cable, etc.

(Category a) Yes, or No, with irrelevant comment or denial of the premiss.
e.g., Yes, he was brave ; Yes, the cable shouldn't be there ; No, he was a show-off.

The judge-reliability for each of the sets of categories was obtained by comparing the gradings given independently by the judges. Expressed as product-moment correlation coefficients (r) they were as follows :

Jane r=·97, SE=·003

Pilot r=·89, SE=·002

IV.—THE POPULATIONS TESTED.

After a preliminary try-out on Oxford 11-year-olds (see appendix), three groups of pupils were tested :

(i) From American schools : New Jersey high, junior high and primary schools.

(ii) From English secondary modern schools.

(iii) From three schools : an Irish boys' and an Irish girls' grammar school and an English comprehensive school.

The American population consisted of 168 pupils, made up of two groups, one of which took the 'Jane' test and the other the 'Pilot' test. The secondary modern population consisted of some 350 boys ranging in age from 11+ to 15+ nearly all of whom took both tests. The comprehensive/grammar school population was composed of 490 boys ranging in age from 11+ to 16+ and 373 girls with an age range of 11+ to 15+. Both boys and girls took both tests.

TABLE 1

DISTRIBUTION OF PUPILS BY AGE IN THE THREE GROUPS.

	Group (i) New Jersey		Group (ii) Secondary Modern	Group (iii) Comp./Grammar	
	'Jane'	'Pilot'		Boys	Girls
16+	3			29	
15+	2	2	79	76	15
14+	25	23	91	101	101
13+	14	9	93	101	102
12+	11	10	77	120	99
11+	13	11	11	63	56
10+	13	12			
9+	9	11			

V.—THE RESULTS.

The tests were presented in group form. This is a possible disadvantage in-so-far as promising leads cannot be followed up as they can by testing each pupil individually. On the other hand, examination of the answers revealed only few replies which it was difficult to categorise.

The results are presented by each test for all three groups of pupils, starting with the 'Jane' test.

TABLE 2

RESULTS FROM 'JANE' TEST.
Ages in months.

Cats	Group (i)					Group (ii) (1)			Group (iii) (2)					
	New Jersey					Sec.-Mod.			Comp./Grammar					
	Mixed					Mixed			Boys			Girls		
		C.A.		M.A.*			C.A.			C.A.			C.A.	
	n	m	s.d.	m	s.d.	n	m	s.d.	n	m	s.d.	n	m	s.d.
cii	20	164·0	12·5	180·5	24·6	90	172·9	13·1	251	168·8	17·0	197	161·8	12·5
ci	10	159·6	13·4	189·5	23·6	43	168·2	13·2	69	167·0	16·4	60	159·0	13·1
bii	25	154·2	20·6	158·5	26·5	86	165·0	10·7	126	159·0	15·6	93	157·6	14·8
bi	32	140·7	23·2	145·4	23·4	100	163·3	12·7	39	149·6	12·6	17	146·5	11·5
a	3	122·7	7·6	116·3	2·4	32	163·6	10·5	5	148·0	4·8	6	152·0	17·5

* M.A.—Verbal reasoning from the Psychological Corporation's Differential Aptitude Test.

TABLE 3

Results from ‘Pilot’ Test.

Cats	Group (i)					Group (ii) (1)			Group (iii) (2)					
	New Jersey					Sec.-Mod.			Comp./Grammar					
	Mixed					Mixed			Boys			Girls		
		C.A.		M.A.			C.A.			C.A.			C.A.	
	n	m	s.d.	m	s d.	n	m	s.d.	n	m	s.d.	n	m	s.d.
cii	24	161·2	15·5	176·2	28·8	42	173·6	11·4	151	172·2	15·4	99	166·3	11·3
bii	34	149·0	21·0	164·3	31·9	197	167·5	13·1	273	163·0	16·6	205	158·3	12·5
bi	10	145·5	21·2	140·5	21·1	59	160·0	12·0	40	148·2	11·3	44	151·9	14·9
a	10	124·5	17·9	134·5	23·7	48	165·4	13·4	12	149·0	8·9	12	152·0	13·7

Obtained by my research students : (1) Mr. B. Whiteman ; (2) Father Paul Andrews.

The standard errors of the differences between the means in the New Jersey results have been calculated for both tests as follows :

TABLE 4 (a)

‘Jane’ Test.

Standard Errors of Differences Between Mean Chronological and Mental Ages.

Chronological Ages				Mental Ages			
Cats.	cii	ci	bii	Cats.	cii	ci	bii
cii				cii			
ci	5·08			ci	9·27		
bii	4·98	5·91		bii	7·64	9·15	
bi	4·97	5·90	5·81	bi	6·88	8·53	6·73

TABLE 4 (b)

‘Pilot’ Test.

Standard Errors of Differences Between Mean Chronological and Mental Ages.

Chronological Ages				Mental Ages			
Cats.	cii	bii	bi	Cats.	cii	bii	bi
bii	4·75			bii	8·03		
bi	7·38	7·61		bi	8·89	8·63	
a	6·45	6·71	8·77	a	9·52	9·28	10·03

In the case of the 'Jane' test results, the chronological age differences are significant at least at the level of twice the standard error between pairs of ages, excepting the differences between categories cii and ci. The 'Pilot' results show similar tendencies. In both tests, the differentiation by mental age is less marked than that by chronological age.

The correlation between the category gradings of the answers of a group of pupils to the 'Jane' and 'Pilot' tests was obtained by Mr. B. A. Whiteman from his secondary modern school pupils. The result was as follows:

$$r_{\mathrm{J.P.}} = \cdot 347, \quad \mathrm{S.E.}_r = \cdot 003.$$

This is a significant correlation since it has been obtained from a single itemed test.

VI.—Discussion.

The purpose of the research is to devise and refine short passages of prose which, in conjunction with the questions asked about them, will enable one to get a clearer view of the growth of children's capacity to make judgments from connected prose. We need to determine whether the progression of the categories by age is matched by a corresponding significant change in the quality of thinking. We do this by looking more closely for the development of logical and explanatory powers.

Omitting Category a with its pre-logical answers, *because I think so, it was her brother who was careless,* we may start with Category bi where we have the stock responses—*Yes, because she didn't watch Teddy . . . , didn't look after Teddy . . . , she let Teddy wander . . .* The one piece of content, or circumstantial evidence, namely that *Teddy wanders . . .*, is seized upon to support the simple unqualified inference that Jane is careless. It is judgment in terms of content which is correlated with the statement implied in the question. Similarly, the pilot's crash is taken to infer that he is not careful. Even though it is limited, this judgment by content consists of a proposition. It is understandable that verbal material should give rise to propositional statements and, as we shall see, the changes in quality of thinking as we pass up to higher categories is marked by increased inference, transformation, complexity and co-ordination of these propositions.

Logically, Category bi answers of the form: *Jane is careless because she let Teddy wander* and *the Pilot is not careful because he crashed* are equivalent to the implications: Teddy wanders (T) implies that Jane is careless (J), and the crash (C) implies that the Pilot is not careful ($\bar{\mathrm{P}}$). This form of argument is the well known *modus ponens* (Suppes and Hill, 1964) argument by which the assertion of the antecedent leads to affirmation of the consequence.

$$\begin{array}{l} \text{T implies J} \\ \text{T} \\ \hline \therefore \text{J} \end{array} \qquad \begin{array}{l} \text{C implies } \bar{\text{P}} \\ \text{C} \\ \hline \therefore \bar{\text{P}} \end{array}$$

Category bii answers produce no basic change in the substance of the propositions of Category bi but an important change in their form. We now have such statements as: *she should have watched Teddy, if she had been careful she would not have . . . ,* and *if the Pilot was careful he would not have crashed the plane*, and *he should have been careful . . .* The argument now takes the form of the well-known *modus tollens* (Suppes and Hill, 1964) by which the thinker considers the denial of the consequences and hence infers the negation of the antecedent.

$$\begin{array}{ll} T \text{ implies } J & C \text{ implies } \bar{P} \\ \bar{J} & \bar{\bar{P}} \\ \hline \therefore \bar{T} & \therefore \bar{C} \end{array}$$

It also marks the use of the conditional and is a more sophisticated form of communication which appears to be separated by a period of 6–12 months from the more simple direct *modus ponens* form.

It would seem to arise as a result of logical maturation but one should not rule out the possibility that in the Category bi answers, the incident (of Teddy wandering or the Pilot's crash) dominates, whereas in the Category bii answers, the trait of carefulness or carelessness dominates over the concrete incident. In both Categories bi and bii, there is, however, no hint yet of including a reference to extenuating conditions and possibilities. This begins to appear in Category ci in ' Jane ' where the stock reply is : *Although Jane was doing her homework she should have watched Teddy*. The claims of homework are now admitted for the first time, and two possiblities appear to be suggested in the stock reply.

First, it is possible that it implies a *practical* action on the part of Jane that she should abandon the homework wholly or partially to look after Teddy. Secondly, there is an important *logical* possibility—that the answer implies that doing homework and watching the child are not incompatible and that wacthing would prevent the child wandering.

There is no Category ci answer in the Pilot passage since, in fact, there are no distracting elements specifically mentioned in the anecdote. To this extent the passages are logically and semantically different. In the Category cii replies we find that the compatibility of homework and watching is strongly questioned, with a consequent shift to justifications for Jane's attention to her homework and for shifting responsibility for the incident to Teddy (young children are very hard to watch, etc.) and even to her mother who should not have placed her in such a dilemma.

The Pilot passage evokes a similar comprehensive judgment involving imagined possibilities to account for the crashing of the plane. The whole transition in thinking from Categories a to cii reveals clear cut and interesting features. These include simple correlations of circumstantial evidence with the final judgment, logical transformations (*modus ponens* to *modus tollens*), combination of propositions, attempts at reconciling propositions and finally recognition of incompatibility and a resulting concentration upon extenuating possibilities.

In Tables 2 and 3 we may note that at mean chronological and mental ages of 13–13½ years, the pupils from all three groups are still judging circumstantially, even though the judgment is expressed in the more sophisticated conditional form (Category bii). Generally, only by 14 years do the pupils show a firm tendency to make more circumspect and comprehensive judgments, by first realising the limited nature of the evidence available and the subsequent invocation of imagined alternative possibilities. This transition from partially descriptive to imagined explanatory responses marks the major change in the quality of the pupil's judgment during mid-adolescence.

This study is, of course, cross-sectional and not longitudinal and this must temper any claims that the results show direct evidence of developmental changes. However, assuming nothing drastically different in the samples, these differences can be described as developmental. The variable age levels associated with the lower category answers (Categories a and bi), highest in the secondary modern group and lowest in the New Jersey high schools group, may be attributed in part to the different age ranges of the samples (see Table 1) and in part to the presumable differences between the I.Qs. of the different groups. Unfortunately, Mental Ages were not immediately available for Groups (ii) and (iii) (see Table 1), but the effects due to differences in the sample age ranges can be eliminated by turning Tables 2 and 3 inside out, as it were, to compute the mean category level of answers produced at different chronological age levels. These follow in Table 5.

TABLE 5

AVERAGE ' JANE ' CATEGORIES PER CHRONOLOGICAL AGE GROUP.

C.A.	New Jersey Boys	Secondary Modern Boys	Grammar/Comprehensive	
			Boys	Girls
15+		3·79	4·74	4 07
14+	3·96	3·34	4·25	4·39
13+	3·64	2·80	4·18	4·34
12+	3·14	2·76	3·75	4·20
11+	2·54		3·38	3·55

When we compare the columns of Table 5 we have differences within any one row, say at 13+ between 3·64 for the New Jersey schools, 2·80 for the secondary modern school and 4·18 and 4·34 for the grammar and comprehensive schools, which may be attributed to the differences in I.Q. between the four groups. It seems that the rise in maturity of answers is partly a function of chronological age and partly a function of intelligence.

VII.—SUMMARY OF THE FINDINGS AND INTERPRETATIONS.

Perusal of Tables 2 and 3 shows that irrelevant, premiss-denying and tautological answers a were given by the youngest pupils studied. The first level of logical judgment bi consists of a proposition (*she's careless because*...), circumstantially linking the one implied in the question with the salient feature of the story. To this extent, the judgment, though logical, is partial and circumstantial. Mean ages of around 12+ seem to be associated with such reasoning. At the next level the form of the judgment is changed but without the addition of any new inferences from the test material (*b*) (ii). This corresponds to a change from the *modus ponens* to the *modus tollens* form of argument, the latter appearing at mean ages of 13 – 13+, that is, up to a year or so older than that associated with the former.

Evidence that pupils take account of other factors, in addition to simple circumstantial evidence, appears in answers associated with mean ages 13+ –14. These answers to the ' Jane ' episode showed the first recognition of the claims of homework and appeared to assume either that doing homework was not incompatible with looking after the child, or, that it should be left unfinished.

Lastly, at the most mature level, the incompatible nature of the problem elements, or their insufficiency, is recognised and there is a shift to produce justifications or explanations. Doing homework and watching the child were seen to be incompatible and in the case of the ' Pilot ' it was recognised that there might have been elements outside his control which brought about the crash. The answers at this stage shift to evoked and imagined suggestions to explain the elements of incompatibility and uncertainty. Chronological ages of 14 and mental ages of 15 seem to be associated with this phase.

In brief, first we get a simple judgment that fails to take account of the incompatible competing element, then a *logical* formal transformation of this judgment by inversion, then the introduction of the competing element without recognising its incompatible role (or if this is recognised, the failure to grasp its necessary nature in the story) and lastly the full recognition of the incompatibility and the corresponding shift of emphasis in the answer to looking for other explanations.

VIII.—REFERENCES.

CASE, D., and COLLINSON, J. M. (1962). The development of formal thinking in verbal comprehension. *Brit. J. Educ. Psychol.*, 32, 103-111.

CROWTHER, G. (1960). *Report*: 15 *to* 18. H.M.S O.

DAVIES, G. (1965) *Concrete and Formal Thinking Among a Group of Adolescent Children of Average Ability.* University of Birmingham (unpublished research)

GOLDMAN, R. J. (1965). The application of Piaget's schema of operational thinking to religious story data by means of the Guttman scalogram. *Brit J. Educ. Psychol.*, 35, 158-170.

LODWICK, A. R. (1958). *An Investigation of the Question Whether the Inferences that Children Draw in Learning History Correspond to the Stages of Mental Development that Piaget Postulates.* University of Birmingham Dip.Ed. dissertation (unpublished).

NEWSOM, J. H. (1963). *Report Half our Future.* H.M.S.O.

PEEL, E A. (1959). Experimental examination of some of Piaget's schemata. *Brit. J. Educ. Psychol.*, 29, 89-103.

PEEL, E. A. (1960). *The Pupil's Thinking.* Oldbourne Press

PEEL, E. A. (1965). Intellectual growth during adolescence. *Ed. Rev.*, 17, 3

RHYS, W. T. (1965). *The Growth of Logical Thought in the Adolescent With Reference to the Teaching of Geography in the Secondary School* University of Birmingham (unpublished research).

STONES, S. K. (1965). *An Analysis of the Growth of Adolescent Thinking in Relation to their Comprehension of School History Material.* University of Birmingham Dip. Child. Psychol. dissertation (unpublished).

SUPPES, P., and HILL, S. (1964). *First Course in Mathematical Logic.* Blaisdell.

IX.—APPENDIX.

The ' Jane ' passage was first tried out as a comprehension question in the Oxford City 11+ examination in 1961.

The passage and questions read as follows :

Jane is a very clever 15-year-old girl who is preparing for her ' O ' level examinations. She has homework every evening and she always works at this diligently, but she is a great help to her mother in looking after the younger children and doing errands and jobs about the house.

One evening, as Jane was doing her homework, her mother asked her to look after her younger brother, Teddy, whilst she went out. Teddy wandered from the living room into the kitchen, got hold of a pot of jam, ate a lot of it, covered his clothes with it and spilled it on to some clean, dry washing in a basket so that it was ruined.

Now answer the following questions ·

(1) Was Jane a careless person ? ....

(2) Why do you think so ? ....

It will be seen that the above passage permits of two plausible and alternative concrete circumstantial responses.

(*a*) No, because she is a great help . . .

(*b*) Yes, because she let Teddy wander . . .

The results show that both responses came up very generally and both are associated with the same level of verbal intelligence.

As the 11+ population had a limited age range of one year, 10 · 4—11 : 4, the verbal I.Q. was used to obtain a mean measure of intellectual competence for each of the answer categories.

A sample was analysed of the answers of some 280 boys, not constituting a random sample* since the least able and slowest working pupils were not able to attempt the question as it came at the end of the English paper.

The results of the analysis were as follows :

Qualities of the answers	New Jersey Category Number	n	Mean Verbal I.Q.
Account taken of extenuating circumstances	cii+ci	32	117·0
Yes, because she let Teddy wander .	bi+bii	116	109·1
No, because she works diligently, helps her mother	bi+bii	76	109·7
No, because she's a clever girl doing her ' O ' level	a+	42	102·4
Irrelevance and misunderstanding	a	15	90·8

The analysis was not as detailed and refined as that devised later for the New Jersey study.

The change in I.Q. level from Category to Category supports the findings from the improved test material which were based on an amended version of the passage to cut out the concrete circumstantial response, *No, because she's a great help* . . . , so making it possible to identify circumstantial thinking by one class of answer only.

---

* It also contained a disproportionately high number of pupils from sophisticated and upper-middle-class homes.

(*Manuscript received* 21*st May*, 1965)

DESIRABILITY RESPONSE SET IN CHILDREN

By SYBIL B. G. EYSENCK, I. A. SYED AND H. J. EYSENCK
(Institute of Psychiatry, University of London)

SUMMARY. The junior form of the Eysenck Personality Inventory, measuring extraversion, neuroticism and tendency to lie, was administered to two groups of children totalling 575, both boys and girls, ranging in age from 11–15. Children in the control group were tested and retested under identical instructions ; children in the experimental group were tested under standard instructions but retested under instructions to 'fake good.' As predicted there were highly significant changes in the experimental group towards greater lying and less neuroticism ; there were no changes in extraversion as compared to the control group. Test/retest reliability in the experimental, as compared with the control group, were slightly lower for extraversion, considerably lower for neuroticism and very much lower for the lie scale. The results were interpreted as giving evidence that social desirability response set was unlikely to have accounted for more than a small portion of the variance in the original test scores.

I.—INTRODUCTION.

THERE has been much argument in recent years about the possibility of response sets or response styles affecting results obtained by use of questionnaires. Rorer (1965) has discussed this evidence at great length and has come to the conclusion that " the data accumulated to date must be interpreted as indicating that response styles are of no more than trivial importance in determining responses to personality, interest and attitude inventories." This conclusion is very similar to that arrived at independently with respect to the M.P.I. and E.P.I. by Eysenck and Eysenck (1963a, 1963b, 1964), and by Eysenck (1962). The recent publication of a junior form of the E.P.I. (Eysenck, 1965) suggested the desirability of investigating the possibility that response sets might be more powerful in the case of children whose responses do not seem to have been studied in detail in this connection. The present study is, therefore, concerned with a test of the hypothesis that the so-called ' desirability ' response set is of little importance in connection with the answers given by children on the Junior E.P.I.

This questionnaire contains twenty-four questions relating to extraversion, twenty-four questions relating to neuroticism and twelve questions which form a lie scale, the purpose of which is to detect faking on the part of the subjects. The scales are reliable and norms are available for quite large age and sex groups. Full information on the test is given in the manual (Eysenck, 1965).

For the purpose of this test ten groups of children altogether were tested, divided into age groups (11–15 years, inclusive), as well as into sex groupse. Two-hundred-and-eighty-nine children were given the test and were then retested within one month, instructions on both occasions being identical and as suggested by the manual. Two-hundred-and-eighty-six children were given the first test in the same way but were then retested after one month with instructions to ' fake good ' ; the specific words used were that they " should try to put themselves in the best light." In essence, therefore, this experiment was a repetition of the one carried out by Eysenck and Eysenck (1963a) on adults, and the underlying idea was that if the original test data had been strongly influenced by social desirability response set, then the deliberate adoption of such a response set should not lead to any great change in scores. If, on the other hand, the original responses had been made without much regard to social desirability, then we would expect considerable changes in the scores of the experimental group, as compared with the control group, particularly in relation to the neuroticism and lie scales. It was not expected that the extraversion scale would be much affected as neither extraversion nor introversion is as obviously more socially desirable as is a low score on neuroticism or a high score on the lie scale.

The detailed results of the experiment are given in Tables 1 and 2. Table 1 shows the actual scores and the changes of the various groups on the three scales, as well as the test/retest correlations under straightforward test/retest conditions. Table 2 gives the same information under experimental conditions. Some of the groups are too small to make their results very convincing but the overall results are quite clear cut. With respect to extraversion, the changes in the experimental group are almost exactly as large as those in the control group, demonstrating that, as in the case of the adult sample, extraversion is not to any reasonable degree subject to conscious faking. When we turn to neuroticism the position is quite different. Changes in the control group average around ·5, whereas in the experimental group they are some six or eight times as high. This difference is very highly significant statistically. It is interesting to note that the changes in the experimental group increase with age from 11 through 12 to 13 ; there is no further increase after 13. This is not an unexpected result, of course, as one might expect increasing age to bring greater sophistication with respect to knowledge of what is and is not socially desirable.

TABLE 1

CONTROL GROUP.

			M			M			M			Test Retest Correlations		
	Age	n	E_1	E_2	E_2-E_1	N_1	N_2	N_1-N_2	L_1	L_2	L_2-L_1	E	N	L
GIRLS	11	38	17·474	17·789	·315	12·658	11·974	·684	3·921	3·921	0	·690	·628	·575
	12	32	18·656	18·906	·250	10·500	9·281	1·219	3·312	3·000	—·312	·786	·804	·63
	13	36	17·389	18·028	·639	11·778	10·889	·889	3·778	3·444	—·334	·715	·748	·780
	14	27	17·296	17·926	·630	12·148	11·926	·222	3·444	3·444	0	·828	·827	·827
	15	7	16·143	17·143	1·000	13·000	12·143	·857	3·571	3·000	—·571	·902	·891	·811
BOYS	11	42	17·643	17·429	—·214	12·286	10·952	1·334	4·357	4·238	—·119	·694	·670	·654
	12	34	17·824	18·176	·352	11·500	10·559	·941	3·529	3·176	—·353	·603	·755	·782
	13	30	18·100	17·933	—·167	11·100	10·667	·433	3·067	3·467	·400	·502	·740	·810
	14	25	17·760	18·840	1·080	9·560	10·760	–1·200	2·120	1·440	—·680	·617	·727	·61
	15	18	18·111	18·556	·445	8·500	7·444	1·056	2·444	2·000	—·444	·821	·861	·78
BOYS AND GIRLS	11	80	17·563	17·600	·037	12·463	11·437	1·026	4·150	4·087	—·063	·688	·655	·627
	12	66	18·227	18·530	·303	11·015	9·939	1·076	3·424	3·091	—·333	·680	·774	·71
	13	66	17·712	17·985	·273	11·470	10·788	·682	3·455	3·454	—·001	·595	·743	·78
	14	52	17·519	18·365	·846	10·904	11·365	—·461	2·433	2·481	·048	·737	·787	·797
	15	25	17·560	18·160	·600	9·760	8·760	1·000	2·760	2·280	—·480	·841	·872	·79

TABLE 2

EXPERIMENTAL GROUP.

				M			M			M		Test Retest Correlations		
	Age	n	E_1	E_2	E_2-E_1	N_1	N_2	N_1-N_2	L_1	L_2	L_2-L_1	E	N	L
GIRLS	11	18	17·111	18·111	1·000	11·500	8·278	3·222	5·222	6·944	1·722	·742	·529	·681
	12	44	18·727	18·864	·137	12 364	9·955	2·409	2·818	5·318	2·500	·573	·692	·210
	13	43	18·070	19·023	·953	12·349	7·047	5·302	2·837	7·372	4·535	·875	·562	·049
	14	54	17 630	19·037	1·407	14·278	8·926	5·352	2·407	7·241	4·834	·415	·597	·093
	15	30	17·167	17·600	·433	14·733	10·200	4·533	2·200	5·767	3·567	·751	·524	– ·008
BOYS	11	17	18·176	18·765	·589	10·000	10·000	0	3·471	3·647	·176	·362	·804	·552
	12	28	18·893	18·893	0	11·250	8·071	3·179	3·071	6·536	3·465	·598	·486	– ·332
	13	34	18·235	19·529	1·294	10·647	6·324	4·323	2·618	6·941	4·323	·698	·675	·282
	14	9	17·556	18·556	1·000	9·667	5·000	4·667	2·111	8·444	6·333	– ·560	·829	·010
	15	9	18·889	17 889	–1·000	11·333	6·444	4·889	1·111	7·667	6·556	·225	·690	·134
BOYS AND GIRLS	11	35	17·628	18 429	·801	10·771	9·114	1·657	4·372	5·343	0·971	·559	·644	·691
	12	72	18·792	18 875	·083	11·931	9·222	2·709	2·916	5·792	2·876	·563	·621	– ·020
	13	77	18·143	19·246	1·103	11·597	6·728	4·869	2·740	7·182	4·442	·782	·604	·150
	14	63	17·619	18·968	1·349	13·619	8·365	5·254	2 365	7·413	5 048	·348	·643	·075
	15	39	17 564	17·667	·103	13·948	9·333	4·615	1·949	6 205	4·256	697	·596	– ·037

TABLE 3

Control Group					Experimental Group				
	E	N	L	n		E	N	L	n
Boys (11–15) years	·640	·726	·743	149	Boys (11–15) years	·509	·632	–·021	97
Girls (11–15) years	·754	·759	·709	140	Girls (11–15) years	·635	·516	·151	189
Total Age and Sex .	·691	·741	·728	289	Total Age and Sex	·596	·608	·093	286

Results on the lie scale are even more impressive than those on the neuroticism scale. For the control group there is a slight tendency for the children to lie less on repetition, but for the experimental group there is a tremendous increase in lying, which again is very highly significant statistically. This increase is almost identical with that shown on the neuroticism scale but is, of course, achieved with a scale having only half as many items and must thus be regarded as being something like twice as large. Again, there is an increase in degree of change from 11 through 12 to 13, and a failure for this increase to continue beyond the age of 13.

The test/retest reliability for both conditions are given in Tables 1 and 2; they have also been calculated for the total group, boys and girls separately, and also boys and girls together; these figures are given in Table 3. It will be seen that reliabilities for the extraversion scale are somewhat lower under experimental as compared to control conditions; those for neuroticism show a further lowering, while those for the lie scale become completely insignificant under experimental conditions. These results are in good accord with the changes in means reported above; although it is, of course, possible for a considerable shift in mean to occur without any change in test/retest correlation, this is unlikely to happen in practice because the added factor of 'faking good' instructions is likely to affect different children differently, and therefore, to lower test/restest correlations.

Our results will be seen to be very similar to those of the adults reported by Eysenck and Eysenck (1963a) in that much greater changes were found for the neuroticism and lie scales than for the extraversion scale. These changes were in the predicted direction, i.e., towards greater lying and less neuroticism and their size indicates that desirability response set is unlikely to have played a large part in the causation of the original scores.

ACKNOWLEDGEMENTS.—We wish to acknowledge the support of the Bethlem Royal and Maudsley Hospital Research Fund. We are also indebted to Mr. K. A. Hooton, Headmaster of South Hackney School, whose co-operation made this study possible, and to Mr. A. Hendrickson, who programmed the data for computer analysis.

II.—REFERENCES.

EYSENCK, H. J. (1962) Response set, authoritarianism and personality questionnaires *Brit. J. Soc. Clin. Psychol.*, 1, 20-24.

EYSENCK, S. B. G. (1965). *Manual of the Junior Eysenck Personality Inventory*. London: University of London Press.

EYSENCK, S. B. G., and EYSENCK, H. J. (1963a). An experimental investigation of 'desirability' response set in a personality questionnaire. *Life Sci*, 2, 343-355.

EYSENCK, S. B G, and EYSENCK, H J. (1963b). Acquiescence response set in personality questionnaires. *Life Sci.*, 2, 144-147.

EYSENCK, S. B. G., and EYSENCK, H. J. (1964). 'Acquiescence' response set in personality inventory items. *Psychol. Rep*, 14, 513-514.

RORER, L. A (1965). The great response-style myth. *Psychol. Bull.*, 63, 129-156.

(*Manuscript received 6th May*, 1965)

RESEARCH NOTES

PERSONALITY FACTORS AND ACADEMIC ATTAINMENT IN JUNIOR SCHOOL CHILDREN

BY R. D. SAVAGE
(*University of Newcastle-upon-Tyne*)

INTRODUCTION AND METHOD.

This report concerns the relationship between extraversion and neuroticism and the intellectual level and school attainment of junior school children.

A sample of ninety-three school children (average age 7 years 11 months, S.D. 6 months) of both sexes, was investigated in two testing sessions. In the first, the Otis Quick Scoring Mental Ability Test, Alpha Form A or Form B (1937) randomly and the Eysenck Personality Inventory (1965) was given. In the second session, either the Schonell Essential Mechanical Arithmetic Tests (Form A) or the Essential Problem Arithmetic (1965) at random and the Watts Sentence Reading Test 1 (1965) were administered. For the group administration of the E.P.I., the author read each question aloud, the pupils followed, then circled the 'Yes' or 'No' as appropriate.

RESULTS AND DISCUSSION.

The major import of this work derives from the intercorrelational data.

TABLE 1

THE INTERCORRELATIONS OF THE MEASURES OBTAINED.

	I.Q.	A.Q.	R.Q.	E	N	L
Intelligence Quotient		·520	·548	·267	—·023	—·026
Arithmetic Quotient .. .			·514	·235	·082	—·249
Reading Quotient				·191	—·217	·066
Extraversion					·068	—·058
Neuroticism						—·419
Lie Scale						

In the first place, it will be seen that there is a significant relation between extraversion and I.Q. (·267), between extraversion and A.Q. (·235), whilst the extraversion—R.Q. (·191) relation fails to reach the 5 per cent. significant level by only ·009. It would appear, therefore, that high extraversion is related to a brighter intellectual level and higher academic attainment in these children. This finding is of some interest as the previous work with students suggested the relationship was in the opposite direction (Furneaux 1957, Gibbins and Savage 1965, Lynn 1959, Lynn and Gordon 1961, and Savage 1962).

The relationship between neuroticism and academic attainment is also of interest in this sample. The neuroticism scores were significantly related to R.Q. (—·217) in a negative direction, but were not significantly related to intelligence or arithmetic. It would appear that the relationships between neuroticism and academic success is not as striking in these children as it was found to be in previous studies of student populations. The explanation of this may well be found in the inverted U-shape relationship between neuroticism and performance (Savage 1962, Savage and Eysenck 1964). These results suggest that the I.Q. and A.Q. test situations did not produce untoward anxiety in this sample, though the reading test situation may well have done so, as the more neurotic children read at a lower level (r= —·217).

The age-neuroticism-lie scale relationships also have interesting aspects : there was a negative relationship (—·419) between neuroticism and lie scale score and a positive correlation between neuroticism and age (·420). This would suggest, unlike the results of Callard and Goodfellow (1962), that the older children score, significantly higher on neuroticism scores, though there was no relationship (·047) between age and extraversion scores. On the other hand, it would appear that the more neurotic the children were, the less likely they were to give lie scale profiles (r= —·419). The lie scale was also significantly related in a negative direction to age (—·536) and it would appear, therefore, that younger children give more lie scale results.

Finally, it may be noted that there is no relationship between neuroticism and extraversion (r= —·068) and they may perhaps be regarded as orthogonal factors as Eysenck has postulated (1957).

REFERENCES.

CALLARD, M PAULINE, and GOODFELLOW, CHRISTINE L. (1962). Neuroticism and extraversion in schoolboys as measured by the Junior Maudsley Personality Inventory. *Brit J. Educ. Psychol.*, 32, 241-250.

EYSENCK, H. J. (1957). *Dynamics of Anxiety and Hysteria.* London . Routledge and Kegan Paul.

EYSENCK, H. J., and EYSENCK, S. B. G. (1965). *Junior Eysenck Personality Inventory.* London University Press.

FURNEAUX, W. D. (1957). *Report to Imperial College of Science and Technology.*

GIBBINS, K. C , and SAVAGE, R. D. (1965). Intelligence, study habits and personality factors in academic success. *Durham Research Review, University of Durham*, 5, 16, 8-12.

LYNN, R. (1959). Two personality characteristics related to academic achievement. *Brit. J. Educ. Psychol.*, 29, 213-216.

LYNN, R., and GORDON, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 114-203.

OTIS, A. S. (1937). *Quick Scoring Mental Ability Test—Alpha Form A and B* London : George G. Harrap and Co.

SAVAGE, R. D. (1962). Personality factors and academic performance. *Brit J. Educ. Psychol.*, 32, 251-253.

SAVAGE, R. D. (1965). The measurement of learning potential in cerebral palsied children. *Teaching the Cerebral Palsied Child.*, Ed. J. Loring. Heinemann.

SAVAGE, R. D., and EYSENCK, H. J. (1964). The definition and measurement of emotionality in Eysenck, H. J. (Ed.). *Experiments in Motivation.* London : Pergamon Press, chapt. 12, 292-314.

SCHONELL, F., and SCHONELL, ELEANOR (1965). *Essentials of Diagnostic and Attainment Testing.* Oliver and Boyd.

WATTS, A. F. (1965). *Sentence Reading, Test I.* London : Nat. Found. Educ. Res.

(*Manuscript received 4th October*, 1965)

EXTRAVERSION, NEUROTICISM AND ACADEMIC PERFORMANCE AMONG GHANAIAN UNIVERSITY STUDENTS

BY PAUL KLINE

(*Department of Education, University of Manchester*)*

SUMMARY. Eysenck's Personality Inventory (E.P.I., 1964) was given to first-year students at the University College of Science Education, Cape Coast, Ghana. Their scores on this test were related to the only common academic measure among them—the intermediate examination in arts. Extraversion was found to be correlated significantly and negatively with success in this examination while neuroticism, though also negatively related, failed to reach significance.

*This study was carried out while the author was Visiting Lecturer in Psychology at the University College of Science Education, Cape Coast, Ghana.

I.—INTRODUCTION.

Several recent studies, Furneaux (1957), Broadbent (1958), Lynn (1959), Bendig (1960), Lynn and Gordon (1961) and Savage (1962) have confirmed the relationship predicted from Eysenckian theory between introversion and academic success among students, though findings concerning neuroticism have been less consistent. The present study was carried out to see whether these personality factors had a similar influence in a culture pattern very different from that of the quoted investigations.

II.—PROCEDURE.

The E.P.I. form A was administered to 110 Ghanaian students at the University College of Science Education, Cape Coast. The representative nature of the sample and the validity of the E.P.I. (and the Cattell 16 P.F. Test) within this population has been demonstrated in a previous study (Kline, 1965). Forty-nine of these students had taken a common examination of the College in 1964 (none of the others had any common academic qualifications which would have have made correlational study possible) and their neuroticism and extraversion scores on the E.P.I. were correlated with the average score obtained on the intermediate examination.

III.—RESULTS.

Mean Score (E) 9·18	S.D 4.49
Mean Score (N) 11.12	S.D. 5·29

Correlations (N=49)

N with Score in Examination r —·229 (N.S.).

E with Score in Examination r —·55 (sig. at ·01 level).

IV.—DISCUSSION OF RESULTS.

It must be noted that the scores of this smaller sample on neuroticism and extraversion were not significantly different from the scores of the remainder of the full sample (110) It can be argued, therefore, that this is a representative, though small, selection of Ghanaian students.

Neuroticism was not significantly related to academic performance in this sample—a finding which supports the work of Furneaux (1957), Lynn (1959) and Bendig (1960). As with these studies a low and non-significant negative correlation was found. Savage (1962) presents evidence for a U-shaped relationship between neuroticism and performance. To investigate this hypothesis, the group was divided into three (high, medium and low academic performance) and an analysis of variance performed on the neuroticism scores, but no significant mean differences were obtained. Thus, no evidence for this U-shaped relationship was discovered. Introversion was strongly related to academic success in this sample, thereby supporting the findings of most investigators in this field : Furneaux (1957), Broadbent (1958), Lynn (1959), Bendig (1960), Lynn and Gordon (1961) and Savage (1962).

From this it appears that even in a culture pattern quite different from that of Western Europe or America, academic success is related to and influenced by the same personality variables. This must be interpreted as evidence in favour of the claim of Eysenck that extraversion and introversion are basically genetically determined, rather than being the products of a particular kind of environment.

V.—REFERENCES.

BENDIG, A. W (1960). Extraversion, neuroticism and student achievement in introductory psychology. *J. Educ. Res.*, 53, 263-67.

BROADBENT, D. E. (1958). *Perception and Communication.* London · Pergamon Press.

FURNEAUX, W. D. (1957). *Report to Imperial College of Science and Technology.*

LYNN, R. (1959). Two personality characteristics related to academic performance. *Brit. J. Educ. Psychol.*, 31, 3, 213-16.

LYNN, R., and GORDON, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 114-203.

KLINE, P. The Use of the Cattell 16 P.F. Test and Eysenck's E.P.I. With a Literate Population in Ghana. (To be published).
Literate Population in Ghana. (In press.)

SAVAGE, R. D. (1962). Personality factors and academic performance. *Brit. J. Educ. Psychol.*, 32, 251-53.

(*Manuscript received 19th August,* 1965)

A NOTE ON SEASON OF BIRTH AND INTELLIGENCE

BY A. B. OJHA, R. P. KELVIN AND C. J. LUCAS
(*University College, London*)

Recent interest in the relationship between season of birth and intelligence has provided data leading to conflicting conclusions. For example, McEwan (1965) and Orme (1962, 1963) claim that Summer (May-July) and Autumn (August-October) born children tend to be more intelligent than Winter (November-January) and Spring (February-April) born children, but Crookes (1963) and Davies (1964). found no such relationship.

The data presented here have been taken from a sample survey* of the 1960 intake of students to University College, London They provide no support for the finding that intelligence and season of birth are related (see Table 1).

TABLE 1

RELATION BETWEEN A.H.5. SCORES AND SEASON OF BIRTH.

(n=145)

A.H.5 Scores	Season of Birth			
	Summer (N=36)	Autumn (N=38)	Winter (N=37)	Spring (N=34)
Mean	37·08	40·39	37·83	39·85
Men	39·89 (19)	41·18 (17)	40·82 (17)	41·23 (17)
Women	33·94 (17)	39·28 (21)	35·30 (20)	38·47 (17)
Part I Mean	18·94	20·14	16·41	19·00
Men	19·25	21·44	17·86	20·11
Women	18·56	18·65	16·96	17·94
Part II Mean	19·47	20·71	18·81	20·54
Men	20·35	21·55	20·00	21·94
Women	18·37	19·94	18·22	19·21

---

* A project supported by the Medical Research Council.

Seventy men and seventy-five women in our sample completed the A.H.5. (Heim, 1956) in the second term of their first year. Season of birth was taken from the date of birth recorded on a social questionnaire completed at approximately the same time.

On A.H.5, there was significant difference between men and women (men, mean=41·10, S.D. 8·35 : women, mean=37·00, S.D. 8·30 ; t=2·81 < ·01). There was no difference in the numbers of men and women born in different seasons—a finding in agreement with the Registrar-General's Statistical Review. In view of sex difference in intelligence, results of men and women were analysed separately.

No significant relation was found between intelligence and season of birth. To ensure that significant differences were not obscured by combining scores of both verbal and non-verbal parts of A.H 5, the scores for parts 1 and 2 were analysed independently. There was no significant difference. The median scores for each part of the test for both men and women were then found and the groups were classified into two : (1) with scores falling on or above the median, and (2) with scores falling below the median. Again, no relation was found between intelligence and season of birth. Sex differences on A H.5. were consistent on both parts of the test (reflecting a very high correlation between parts 1 and 2) and season of birth.

In conclusion, we found no evidence of a relationship between season of birth and intelligence.

REFERENCES

CROOKES, T. G. (1963). A note on intelligence and season of birth. *Brit. J. Med. Psychol.*, 36, 355-6.

DAVIES, A. D. M. (1964). Season of birth, intelligence and personality measures. *Brit. J Psychol.*, 55, 475-6.

HEIM, A. W. (1956). *Manual for the Group Test of High-Grade Intelligence A.H.5.* National Foundation for Educational Research.

MCEWAN, P. J. M. (1965). Climate and intelligence. *Brit. J. Soc. Clin. Psychol.*, 4, 8-13.

ORME, J. E. (1962). Intelligence and season of birth. *Brit. J. Med. Psychol.*, 35, 233-4.

ORME, J. E. (1963). Intelligence, season of birth and climatic temperature. *Brit. J. Psychol.*, 54, 273-6.

(*Manuscript received 15th July, 1965*)

THE SECULAR TREND IN PUBERTY : HAS STABILITY BEEN ACHIEVED ?

BY PAMELA K. POPPLETON (*Department of Education*) AND
P. E. BROWN (*Department of Preventive Medicine and Public Health, University of Sheffield*)

SUMMARY An investigation is described in which data concerning the age of menarche were obtained from a sample of girls in three West Riding schools. Every care was taken to ensure the accurate reporting of this event. When the figures were compared with those obtained earlier from a closely similar group, a discrepancy of six months in the median age was noted. This is explained in terms of an artifact resulting from the method of reporting the event. When the more recent figures were compared with data obtained from other surveys conducted in the last sixteen years, no consistent secular trend was revealed.

INTRODUCTION.

The existence of a trend towards earlier physical maturity (commonly known as the secular trend), has been described and its progress recorded by surveys conducted in many countries between 1840 and 1960. The evidence is summarised by Tanner

(1961, 1962) and has been widely quoted by writers on medical, social and educational problems. Several indices of maturity have been used in these surveys, the most readily obtainable being the age at which girls commence menstruation. Sometimes this event has been reported at the time of its occurrence, as in the Norwegian and French surveys mentioned later, and sometimes it has been recalled at a later age. The inverval of recall has varied from a year or less after the event (Nisbet and Illesley, 1963) to what is almost a lifetime (Brown, 1966). The accuracy of recall in this type of investigation has several times been questioned, but only one attempt has been made to examine systematically the nature and extent of the error (Livson and McNeill, 1962). It is the purpose of this paper to describe a recent investigation among girls in the West Riding of Yorkshire which was undertaken to examine the hypothesis that errors of recall existed and that they might be related to social and psychological factors. The investigation revealed certain discrepancies by comparison with other surveys, the examination of which leads to some tentative conclusions regarding the secular trend.

THE INVESTIGATION AND ITS BACKGROUND.

In 1962, a Thyroid survey was conducted in the West Riding. Six-hundred-and-eighty-five girls between the ages of 15 and 19, who were still attending school were examined and, amongst other things, were asked to state the year and month in which their periods started. From this information the age of menarche was calculated for each girl and a frequency distribution was drawn up which, when compared with earlier ones, showed evidence of a continued secular trend. The mean age of menarche appeared to have fallen from 15 years 3 months in 1845, to 12 years 9 months in 1962, and the phenomenon is discussed in an article by one of the authors (Brown, 1966). An interesting feature noted in the 1962 figures was a ' rounding off ' of the dates given. It was observed that the most preferred months mentioned by the girls tended to cluster round the birthdate. Relatively few girls appeared to start their periods in the mid-year between birthdays.* It seemed that these differences and irregularities might be accounted for by faulty recall on the part of some of the girls, and that recall might be influenced by psychological and social factors.

Consequently, a further investigation was undertaken in 1965 which was designed to produce as accurate and reliable data as possible from a substantial sample of girls Three of the schools involved in the 1962 survey expressed their willingness to co-operate and a simple questionnaire was prepared for the girls to complete under the supervision of the school matrons.† The schools were a grammar school, a secondary modern school (now one unit) and two comprehensive schools. All girls in the first four years of secondary schooling who were then between the ages of 11½ and 15½ were asked to indicate whether their periods had started or not and, if so, the school term or holiday during which this event had occurred. The questionnaire was first administered after the Christmas holiday in January, 1965, and then again following the Easter and Summer holidays. After these three administrations, records for the fourth-year girls were practically complete, and they are currently being completed for the first, second and third-year girls. It was particularly emphasised that during the administration, the girls must not be helped to remember and that if a girl was uncertain she was to endorse ' Not sure.' By taking the mid-point of each term and holiday period, an approximate age of menarche was obtained for each girl, Bearing in mind the uncertainty which surrounds this event, accuracy to within a month is probably as much as can be expected with errors distributed at random in either direction.

Thus, two sets of data were obtained within a period of three years from girls in the same area, attending the same schools and as similar in every respect as possible. The differences were that in 1962 girls were questioned between the ages of 15 and 19, whereas in 1965 they were between the ages of 11½ and 15½. In 1962, they were asked to recall the year and month of menstruation which was recorded by the school nurse and, in 1965, they were asked to recall on the basis of a prestructured questionnaire administered under standardised supervision.

---

* The month by month distribution was not uniform. For the total number of 734 girls, the deviation from the expected monthly value of 61 gave a highly significant value for χ^2 (31·8, df. 11, p<0·01).

† Since 1960, schools over a certain size (1,000 pupils) have been entitled to the services of a full-time school matron in the West Riding.

RESULTS AND DISCUSSION OF RESULTS.

TABLE 1

CUMULATIVE PERCENTILE DISTRIBUTION OF GIRLS COMMENCING THEIR PERIODS BY A GIVEN AGE. (Three West Riding Schools).

	1962	1965
	N=685	N=343
Before 12 : 0	20·0	12·2
12 : 0—12 : 5		23·9
12 : 6—12 : 11	55·3	35·3
13 : 0—13 : 5		54·5
13 : 6—13 : 11	84·9	69·9
14 : 0—14 : 5		84·4
14 : 6—14 : 11	96·9	933·
15 : 0—15 : 5		95·0
Over 15 : 6	100·0	100·0
Median Age	12 : 10	13 : 4
Semi-interquartile range	9 months	10 months.

The frequency distributions shown in Table 1 are those for all girls questioned in 1962, and for those girls between the ages of 14½ and 15½ questioned in 1965. These girls were in their fourth year of school and are the only ones for whom a complete record is yet available. Twenty-four girls who had started their periods but were unsure of the date, were omitted from the 1965 figures. They formed 6·5 per cent. of the total of 367.

It will be seen that there are large discrepancies between these two distributions and that the median age in 1962 was six months earlier than in 1965. If, however, the reader now follows the lines which have been drawn in, he will find a remarkable correspondence between the linked figures, those in the second column being the percentages obtained at an age six months later than in the 1962 distribution. How can this phenomenon be accounted for ?

In 1962, Livson and McNeill published the results of a unique longitudinal study in which a small sample of women for whom the age of menarche had been recorded seventeen years earlier, were asked to recall this age. Forty-three women who were an unbiased sample from a larger group were questioned and gave a mean recalled age of 12·3 years compared with a mean actual age of 12·8 years. This is a precise error of underestimation of six months. The frequency distribution of recalled menarcheal ages showed, according to the authors, " a serrated contour with steep valleys at half-year points and sharp peaks at full-year points, making very apparent a tendency for these women to report menarcheal age in whole years." This is similar to the phenomenon described in the 1962 West Riding data, and we agree with Livson and McNeill that the error is one of report rather than recall and that it is highly dependent upon the procedure used. They go on to say, " Thus, a recall of menarche as having occurred ' when I was 12,' almost certainly refers to the interval whose limits are 12·0 and 13·0 and whose midpoint is 12·5." In both the cases discussed, the interviewer recorded menarcheal age as interpreted by the girls. Their reasoning followed the lines, ' I was 12 until I was 13 ' and this evidently contributes to the discrepancies observed. In the 1965 data no such peaks and valleys are evident, and it is clear that the girls were able to locate the date by reference to a framework of events (school terms and holidays). It is interesting that, whilst one might expect an error due to recall after a long period of time, the error due to the method of reporting can occur at any time after periods have started.

IMPLICATIONS FOR THE SECULAR TREND.

Frequency distributions such as those in Table 1 have provided the basic data upon which evidence concerning the secular trend is based. Brown (*op. cit.*) compares the distributions from eight such surveys conducted between 1845 and 1962, which appear to offer supporting evidence. A number of additional surveys which have been carried out during the last sixteen years have been located. The reliability of these surveys is more open to scrutiny and it was decided to compare our own figures with those derived from cases in which the methodology is known. We have limited ourselves to surveys conducted in this country or in other Western European countries.

TABLE 2

AGE OF MENARCHE ACCORDING TO VARIOUS SURVEYS (1949—1965).

	1949 Oxford	1950 France	1952 Oslo	1960 Aberdeen	1961 Eng. and Wales	1965 West Riding	
						3rd year	4th year
Age	N= 1,338	N= 75,000	N= 11,618	N= 1,385	N= 1,812	N= 313	N= 343
Before 12 : 0 ..	7·0	12·5	12·3	12·0	15·6	9·9	12·2
12 : 0—12 - 5 ..		23·1		24·62	30·1	21·4	23·9
12 : 6—12 : 11	28·4	44·2	41·2	40·15	42·3	36·7	35·3
13 : —13 : 5 ..		60·0			61·3	56·8	54·5
13 - : 6—13 : 11	63·6	72·7	74·7		77·0	73·5	69·9
14 : 0—14 : 5 ..		80·4			87·5	78·9	84·3
14 : 6—14 : 11	89·6	84·6	93·2		100·0	79·55	93·3
15 : 0—15 : 5 ..							95·0
15 : 6—15 11	95·4						100·0
Above 16 : 0 ..	100·0						
Median	13 : 6	13 : 2	13 : 3	13 : 2	13 : 2	13 : 4	13 : 4
Semi-interquartile range	9 months	12 months	9 months	incomplete	10 months	incomplete	10 months
Method........	School records	Yes-No	Yes-No	Recall 13–16	Recall at 15	Recall at 13+	Recall at 13+

Seven distributions are compared in Table 2, including the, as yet, incomplete distribution for our third-year girls. The methods used are in each case described at the foot of the column as either ' Yes-No ' or ' Recall ' at a given age. ' Yes-No ' indicates a method in which random sample of girls at each half year of age is asked to indicate whether or not menstruation has started. It will be observed that the 1960 (Aberdeen) survey is incomplete. Nisbet and Illesley were mainly concerned with identifying early maturers in their sample. Douglas's 1961 classification groups together all those who had reached puberty after the age of 14 : 6. Both, however, yield a median age though, in the Aberdeen survey it is likely to be a slight underestimate as the figures fall just short of the 50 per cent. mark. Douglas's figures were given in a personal communication (1965). The Oxford figures are a weighted composite of those given for two distributions by Wilson and Sutherland (1949). The ' Yes-No ' method used in the two national surveys, Aubenque (1964) and Kŭl (1953) gives the most reliable results and so a check on the median for the 1965 West Riding sample has been calculated using this method. From the whole sample of 1,454 girls, the proportion who had started their periods at each of the six-monthly intervals was calculated. This gave a median of 13 years 4½ months for which the 95 per cent. confidence limits were 13 years 3 months and 13 years 6 months.

In interpreting these figures one must take into consideration the variations which might be expected to arise in samples of such differing size and representativeness. Certain factors are known to correlate positively with early maturity, e.g., an urban environment (Wilson and Sutherland) and measured intelligence (Nisbet and Illesley, Douglas and Ross). Our own sample, though small, covers the whole range of ability and is drawn from schools in a mixed rural and mining area. Bearing these factors in mind, the results give a picture of remarkable consistency over the period of sixteen years. Tanner (1961) has estimated the extent of the secular trend at " Four months per decade in Western Europe over the period 1840 to 1960. Thus, a girl today may expect to menstruate on average some ten months earlier than did her mother." On the basis of this estimate one would expect to find a decline of just over six months during the period examined, but we observe very little fluctuation in the median age between the limits of 13 : 2 and 13 : 6. More particularly there is no evidence of any systematic downward trend.

Although no comparison with American figures has been attempted, it is worth noting that in an article which also examines the evidence for earlier maturation, McCammon (1965) gives 13.17 years as the mean age for children studied by the Child Research Council in Denver, Colorado. This is the same as the mean age found by this body in the 1930's. McCammon criticises the inadequate data upon which early information was based and concludes : " We need studies in large numbers of girls in which the day, month and year of the menarche are recorded with the same accuracy as the day, month and year of birth," a sentiment with which the authors wholeheartedly agree. No study so far undertaken has satisfied this criterion because no education or medical authority has so far considered this information to be of vital importance, notwithstanding their readiness to cite the secular trend in support of arguments and decisions. In spite of the difficulty of obtaining reliable evidence of this event, it is concluded that, whatever the situation before 1949, since then the age of menarche appears to have remained stable and there is no evidence of a continuing delcine.

ACKNOWLEDGEMENTS.—The authors would like to thank the West Riding Medical and Education authorities for supporting this investigation and the staffs of Wath County Grammar, Wombwell High and Swinton Comprehensive schools for their co-operation.

REFERENCES.

[1] AUBENQUE, M. (1964). Note documentaire sur la puberté et la taille des filles. *Biotypologie*, 3-4, pp. 136-147.

[2] BROWN, P. E. (1966). The age of the Menarche. *J. Prev Soc. Med.*

[3] DOUGLAS, J. W. B , and ROSS, J. M. (1964) Age of puberty related to educational ability, attainment and school leaving age. *J. Child. Psychol., Psychiat*, 5, pp. 185-185.

[4] KÜL, V. (1953). Menarche-allderen hos skolepiper i Oslo ag sammenhangen mellom menarche-alder ag fysisk utvikling. *Statist. Kvartalscheft*, Oslow, 43, pp. 84-88

[5] LIVSON, N., and MCNEILL, D. (1962) The accuracy of recalled age of Menarche. *Human Biology*, 34, pp. 218-221.

[6] MCCAMMON, R. W. (1965) Are boys and girls maturing physically at earlier ages? *Amer. J. Pub. Health*, 55, pp 103-106.

[7] NISBET, J. D., and ILLESLEY, R. (1963) The influence of early puberty on test performance at the age of eleven. *Brit. J. Educ. Psychol.*, xxxiii, pp. 109-106.

[8] TANNER, J. M. (1961). *Education and Physical Growth.* London Univ. Press.

[9] TANNER, J. M. (1962). *Growth at Adolescence.* Oxford : Blackwell.

[10] WILSON, D. C., and SUTHERLAND, I. (1949). The age of the Menarche. *Brit. J. Med.*, 2, pp. 130-132.

A NULL RELATIONSHIP BETWEEN ECOLOGY AND ADULT EDUCATION

BY TERENCE LEE

(*University of St. Andrews*)

One of the decisions to be made in providing buildings and facilities for education is where they should be sited. ' Convenience ' is the main psychological criterion and when attendance is voluntary, as in adult education, this assumes considerable importance. Journeys need to be minimal but without sacrificing the material advantages of centralisation.

The aim of the present study was to determine the effects of distance on enrolment. It arose from previous work where the probability of joining social clubs was found to be a function of their proximity and the common notion of a stable dichotomy of ' joiners ' and ' non-joiners ' was rejected (Lee, 1963). It was felt, however, that the ecology for ' serious ' activities might be different.

PROCEDURE AND RESULTS.

Data were collected from a college providing a wide range of adult education, near the centre of a provincial city (population 85,000). All enrolments (N=424) for one session of the English and Social Studies Department were included. The classes concerned were non-vocational, designed for part-time evening students and included drama, music, languages, literature and other subjects of general cultural interest.

Concentric circles, half a mile apart, were drawn on a map of the city to form eight annuli with the college at centre. Addresses of all enrolees were plotted and expressed as a ratio of total population for each annulus. Population figures were derived from the electoral roll and converted to age 16+ (enrolment age at the college) by a correction figure based on the total population.

Enrolment rates per annulus (Table 1) suggest that enrolment is unaffected

TABLE 1

PROXIMITY AND RATE OF ENROLMENT.

	Distance in miles, less than—								
	1/4	1/2	3/4	1	1·1/4	1·1/2	1·3/4	2 approx	N
Enrolment per 5,000 popn. ..	30	38	32	38	33	36	42	42	
Actual enrolment 'O' ...	22	60·5	57	53	67·5	71	44	49	424
Null hypothesis 'E'	26	57·3	64·2	53·3	72·6	70·8	37·7	41·9	

$$\text{'E'} = \frac{\text{total no of users}}{\text{total population}} \times \text{population in annulus.}$$

Chi-square=4·22. A coefficient of 14·07 would be required for significance at the ·5 per cent level. The small degree of relationship which is shown is in the reverse direction.

by convenience. It seemed possible, however, that the hypothesised effect is nullified by an opposing influence from social class. Thus, middle class people (with more previous education) may be more strongly attracted to enrol, but because they concomitantly live in outer suburbs, the expected fall-off with distance is counterbalanced.

To check this possibility, enrolments were plotted for the thirty-eight polling districts of the city and again expressed as a ratio of population. The rateable values of all dwelling units were obtained from a Land Use Survey on punched cards and means computed for each polling district. When the two most highly rated districts (which show some reversal of the trend) are excluded, r between mean rateable value and enrolment is ·82 ; for the total sample r= ·64.

It was not feasible to control social class in re-testing the main hypothesis, but records were available for *previous education*, probably a more critical variable. It was decided, also, to use a more sensitive index of use than enrolment, i e., *attendance*, and to employ an analysis of variance design.

Students were sampled randomly from the record forms and allotted as appropriate to one of eight distance groups and two categories of previous education (secondary modern or grammar) until each of the sixteen cells contained six Ss.

None of the F ratios was found significant. Attendance is unaffected by distance with previous education held constant. It is also interesting that although previous education strongly effects enrolment, it is not related to attendance, once enrolled.

CONCLUSIONS.

The first implication for urban planning is that adult education facilities can be sited at least five miles apart without loss of potential students. Motivation to participate, though much rarer, must be substantially stronger than for 'light' social activities. The advantages of centralisation in terms of equipment, staff and class sizes, etc. (which in any case are much greater) can be fully exploited.

A second point is that the policy of combining both light and serious activities in a neighbourhood building, which is implicit in the community centre movement, may need re-examination. If serious activities do not respond to convenience, the neighbourhood could hardly be an adequate catchment area. Even if it could attract its territorial proportion, the numbers would scarcely be viable. In fact, however, there would be drastic depletion resulting from the restricted choice of activities that could be offered.

The sense of failure felt when many post-war centres have died or become ' mere social clubs ' might be put down to unrealistic goal-setting.

REFERENCE.

LEE, T. R. (1963). The optimum provision and siting of social clubs. *Durham Res. Rev.*, 4, 53-61.

(*Manuscript received 25th January*, 1965)

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

JUNIOR CHILDREN'S RESPONSES TO THE MURRAY THEMATIC APPERCEPTION TEST AND THE CATTELL PERSONALITY QUESTIONNAIRE

By J. M. MUNDY

(**Abstract of thesis submitted for the degree of Master in Education of the University of Manchester,* 1963)

The purpose of the experiment was three-fold:

(1) To decide which T.A.T. variables occurred most frequently in children's responses.

(2) To discover if there were any sex differences in response and, if so, whether these reflected sex differences in overt behaviour.

(3) To discover if there was a consistent relationship between T.A.T. responses and responses to a personality questionnaire.

ADMINISTRATION AND SCORING OF THE T.A.T.

The subjects were forty-five boys and forty-five girls in their last year in the primary school. Using a method of group administration, written responses were obtained to nine pictures selected from the standard set on the basis of their apparent suitability for children. Stories written about the first picture were discarded and the rest, bearing only index numbers, were scored by the first scorer for 22 Need and 18 Press variables. The scorer used Avon's (1949) version of Murray's definitions of Need and Press, although certain modifications were made where confusion seemed likely to arise. A variable was scored each time it appeared and no account was taken of intensity.

FREQUENCY OF OCCURRENCE OF VARIABLES.

There was wide variation in the extent to which different variables occurred, less than 12 per cent. of children mentioned some, while others were referred to by over 90 per cent. The following table gives information relating to variables most frequently scored. The findings here were somewhat similar to those of Sanford (1943).

NEED		PRESS	
Variable	% of children scoring	Variable	% of children scoring
Need Affiliation	96·67	Press Nurturance	97·78
Need Cognizance	95.56	Press Aggression	84·44
Need Acquisition	88·89	Press Physical-Danger	83·33
Need Achievement	83·3	Press Gift	82·22
Need Succorance	80·00	Press Dominance	81·11

RELIABILITY OF SCORING.

Ninety stories were rescored by two independent scorers and, for each variable where scoring was high enough to permit it, agreement between the scores of the first scorer and each of the two later scorers was checked by means of the chi-squared technique. The reliability of the scoring for 7 Need and 7 Press variables was checked in this way. The scores were so restricted that it was only possible to check agreement in the categories ' scoring ' and ' not scoring ' : chi-squared values ranged from 16·18 to 53·89 and with only 1df this represented a high degreeof association between different scorers.

SEX DIFFERENCES IN T.A.T. RESPONSES.

T.A.T. scores were examined for sex differences in numbers scoring, sex differences in means and sex differences in variability. Scores were converted to ' T ' scores to normalise distribution and corrected for length of response, where necessary, by means of a method suggested by Lindzey and Silverman (1959). There were no significant differences in numbers scoring but there were the following differences in mean score :

Need Succorance	B>G	(·01 level)
Need Acquisition	B>G	(·05 level)
Need Recognition	B>G	(·05 level)
Press Achievement	G>B	(·05 level)

In the case of Need Acquisition, Need Recognition and Press Achievement, these differences were parallel to sex differences thought to exist in overt behaviour. The sex difference in the case of Need Succorance was opposite to the difference thought to exist in overt behaviour.

Boys' scores were found to be more variable than girls' for Need Aggression and it was felt that this might reflect a greater interest in overt Aggression on the part of boys leading in some cases to high scoring and in other cases to low scoring because of attendant anxiety.

COMPARISON OF T.A.T. SCORES WITH SCORES ON A PERSONALITY QUESTIONNAIRE.

Forms A and B of the Cattell C.P.Q. test were administered to the ninety subjects of the experiment. The test was scored for 14 first order factors and the second order factors Extraversion-Introversion and Anxiety-Adjustment. Scores obtained were correlated by means of the Atlas computer with 18 Need and Press variables and the combined variable Need+Press Aggression. Sixteen correlations were found to be statistically significant at the ·05 level and seven at the ·01 level.

In five cases it was found that T.A.T. variables were associated with C.P.Q. factors bearing some resemblance to them in meaning :

Press Loss was associated with factor Q4+.
Need Recognition was associated with factor E+.
Need Affiliation was associated with factor J− and factor O−.
Need Cognizance was associated with Extraversion.

In four cases it was found that T.A.T. variables were associated with C.P.Q. factors contrary to them in meaning :

Need Harm avoidance was associated with factor F+.
Need Succorance was associated with factor A −.
Press Loss was associated with factor F+.
Need Affiliation was associated with factor Q4+.

In other cases correlations were difficult to explain by any discernible correspondence in meaning between T.A.T. variables and associated C.P.Q. factors.

(*Manuscript received* 19*th August*, 1965)

AN INVESTIGATION INTO TWO METHODS OF TEACHING ENGLISH IN MODERN SCHOOLS

By D. COOPER

(Abstract of thesis submitted for the degree of M.Ed. of the University of Manchester)

The purpose of this inquiry was to compare experimentally the efficiency of two methods of teaching English.

The Methods.

The Analytical methods involved the deliberate analysis of English into reading, spelling, usage, vocabulary, punctuation, sentence structure and written composition. Teaching by these methods entailed instruction, exercises and drills in these separate aspects of the subject because it was believed that, as there was improvement in each part, there would be a corresponding improvement in the general English attainment of the children, particularly as reflected in their written composition work.

The Incidental methods involved the treatment of English as a whole subject, emphasis being put on extensive reading and writing. As there was no deliberate analysis of English into separate skills, the use of drills and exercises was avoided, supporters of these methods claiming that the progress of children would be just as great if the children concentrated on extensive reading and writing.

The Questionnaire.

A questionnaire was constructed and sent to 105 secondary modern schools in parts of Cheshire and South-East Lancashire. It was concerned with the teaching of English to first-year classes and was completed by fifty-two schools.

The Experiment.

Lasting for seven months, the experiment was conducted in four secondary modern schools situated in South-East Lancashire, each school providing two first-year classes of children, aged 11·3—12·3 at the beginning of the experiment. Two of the schools normally used Analytical methods, the other two normally used Incidental methods. Four classes, one in each school, were taught by teachers using Analytical methods; the other four classes were taught by different teachers using Incidental methods. In this way it was arranged that two classes using each method were taught by teachers using their normal methods and the other two were taught by teachers using imposed methods. The advantage of this arrangement was that it minimized the effects of two of the major factors often associated with teaching ability—those of experience in the use of, and enthusiasm for, a method.

The Tests.

The tests used were: Verbal and Non-Verbal Test 1, issued by the National Foundation for Educational Research, Graded Word Reading Test and Graded Word Spelling Test Form A by F. J. Schonell, and The Diagnostic English Tests by F. E. Schonell. Each of the English tests was given twice—once at the beginning and once at the end of the experiment—and the results provided the initial and final scores used in the statistical analyses.

The Analyses.

The test data were analysed by the technique of analysis of covariance by means of which the uncontrolled factors of initial ability and, where necessary, intelligence were held constant.

The arrangement of teachers to methods permitted further analyses to show the relative efficiency of teachers using methods to which they were accustomed and about which they were enthusiastic, and of teachers using methods imposed upon them by the conditions of the experiment.

THE RESULTS.

(1) In tests of reading, spelling, usage, vocabulary and written composition, there was no evidence of significant differences in level of achievement between children taught by Analytical methods and similar children taught by Incidental methods, and therefore, both methods were regarded as equally effective.

(2) In tests of punctuation and sentence structure children taught by Analytical methods made significantly greater progress than those taught by Incidental methods. The evidence indicated that, where English was taught by Incidental methods, some supplementary instruction and practice in punctuation and sentence structure would be necessary, if children using the methods were to make as much progress in these mechanical aspects as similar children using Analytical methods.

(3) In all tests, except reading where no significant differences were found between groups, children taught by experienced teachers, who were enthusiastic about their methods, made significantly greater progress than similar children taught by teachers using imposed methods. The evidence available from this part of the experiment offered some support for the beliefs that an enthusiastic and experienced teacher would get good results no matter which method he used, and that, in methods experiments, the teacher variable is an important one, which could produce bias in favour of one of the methods unless the effects of the variable were controlled in some way.

(*Manuscript received* 19*th August*, 1965)

TEACHERS' ASSESSMENTS OF THE AIMS OF TEACHING ENGLISH IN SECONDARY SCHOOLS

BY BRENDA M. CASEY

(*Abstract of thesis submitted for the degree of Master in Education of the University of Manchester*, 1964)

AIM OF THE INVESTIGATION.

The aim of the investigation was to assemble a set of objectives for the teaching of English in secondary schools, to provide an evaluation of them by practising teachers and training college lecturers and to compare the evaluations of teachers in different types of school and of teachers grouped according to age and sex.

METHODS OF INVESTIGATION.

(1) A questionnaire was compiled consisting of sixty objectives taken from books on the teaching of English published during the period 1950-59. This questionnaire was completed by thirty-six teachers in secondary grammar schools, thirty-six teachers in secondary modern schools, and thirty-six teachers in secondary technical schools under the Manchester local education authority, and by 21 lecturers in English at training colleges in the Manchester area training organization. Teachers were asked to supply information as to their age, sex and previous teaching experience and to assess each of the sixty objectives (grouped into six categories) on a nine-point scale. Six teachers agreed to rate the questionnaire a further three times, on each occasion with reference to pupils of a particular age and level of ability.

(2) The data were then analysed in the following ways :

A. (*a*) Two analyses of variance were carried out on the category scores of the male and female groups in different secondary schools and age groups to determine whether significant differences existed between age groups, between teaching groups and between categories.

(*b*) (i) Mean scores for each item were calculated for the total sample and for age, sex and school groupings. These were then used to obtain the rank order of items for each of the groups in each of the six categories of the questionnaire (viz., Writing, Prose Reading, Mechanics of English, Oral Expression, Poetry and Drama).

(ii) Rank order correlations were then computed for each of the groups for the six categories of the questionnaire.

(iii) Tests of significance were applied to the mean scores on each item of the questionnaire in order to determine whether any differences existed between men and women teachers, between the school groups and between the age groups.

B. Each item was reclassified under one of the following three categories : Intellectual, Socio-Emotional and Aesthetic. The median score for each category was calculated for the total sample and for age, sex and school groupings. The scores of the groups for each item were examined in relation to the median scores to see whether new patterns and differences emerged.

C. The scores of the three sets of the questionnaire completed by each of six teachers were individually analysed in order to determine differing patterns of emphasis when the age and ability of pupils were taken into account.

RESULTS OF THE INVESTIGATION.

(1) The reliability of the evaluations was obtained by the test-retest method. Reliability coefficients were computed for each of the sixty objectives, the average correlation being ·79.

(2) Two analyses of variance, one for men teachers, one for women teachers, revealed :

(i) Significant differences at the ·01 level among the six categories of objectives, objectives in the Prose Reading and Oral Expression categories being most favourably received and objectives in the Poetry and Drama categories being least highly received.

(ii) No significant differences amongst the means of the three school groups and amongst the means of the three age groups.

(3) An analysis of the scores for the objectives in the six categories of the questionnaire by the teachers as a whole revealed the following information :

(i) *Objectives in the Writing Category.*—Objectives which referred directly to writing such as, " to encourage the pupil to write creatively in an interesting and lively manner," were the most popular ; those concerned with the preparatory stages of writing were next in order, and objectives which were designed to help the pupil's social and vocational development were least popular.

(ii) *Objectives in the Prose Reading Category.*—Teachers regarded the most important function of reading to be that of giving pleasure to the pupil. They were also favourably disposed towards objectives concerned with developing his powers of comprehension and judgment and widening his experience and general knowledge. Teachers rated less highly objectives seeking to improve the pupil's proficiency in writing and speaking.

(iii) *Objectives in the Mechanics of English Category.*—Teachers considered the greatest value of this aspect of English to be that of helping the pupil to organize and express his thoughts clearly and accurately. Objectives concerned with the development of skills in interpretation and with the supplying of standards which the pupil would apply to his own work were also rated highly but appreciation of the beauty of language structure and facility in adapting language skills to vocational and social situations were considered less essential

(iv) *Objectives in the Oral Expression Category.*—Objectives concerned with ease and fluency of speech in informal situations were rated more highly than those which refer to giving the pupil experience of formal speech situations. The ability to judge the quality of one's own and other people's speech was rated fairly highly. Objectives which stress the importance of increasing the pupil's confidence and motivating him to speak well were also considered important, but teachers were reluctant to recognize the therapeutic value of oral expression.

(v) *Objectives in the Poetry Category.*—Teachers regarded the most important objective to be that of stimulating a love of poetry, and they also recognize the value of poetry in helping to increase the pupil's knowledge and appreciation of language of quality. Objectives concerned with the development of the pupil's confidence and dramatic abilities and with the improvement of his speech and writing through poetry were less well received. 'Transfer of training' objectives, which sought to improve the pupil's memory and powers of concentration through poetry, were rated lowest.

(vi) *Objectives in the Drama Category.*—An objective, "to give the child a love of drama," was rated highest. Teachers were also well-disposed towards objectives concerned with increasing the pupil's knowledge of plays and stimulating a critical interest in them, but they were less inclined to rate highly objectives concerned with the development of interest in literature and related subjects through drama. Objectives concerned with the inculcation of desirable personal and social attributes were received with only a moderate degree of approval.

(4) Variables of age, sex and type of school:

(i) *Age groups.*—There was a very high degree of correspondence among the three age groups in their ranking of the objectives. The youngest age group was in many instances significantly less generous in the rating of objectives than were the other age groups.

(ii) *Men and women teachers.*—There was a considerably high degree of correspondence between the men and women teachers in their ranking of the objectives. Women teachers tended to score more generously than men, particularly on those objectives which are child-centred and liberal in character.

(iii) *School groups.*—The four school groups were in reasonably high agreement over the ranking of objectives. Ranking by training college lecturers was more in agreement with that of secondary modern school teachers than it was with that of other school groups. Secondary modern and secondary technical school teachers were on the whole more in agreement with each other than they were with secondary grammar school teachers. There was a marked tendency for grammar school teachers to rate formal objectives more highly than did the other school groups, and to approve less of vocational objectives.

(5) In general, objectives which stressed aspects of intellectual development and activity were rated more highly than objectives concerned with the social, emotional and aesthetic development of the pupil.

(6) Analysis of the rating of six teachers who made separate evaluations for children of different ages and abilities revealed only slight differences in the ratings for most of the objectives but fairly substantial differences in the ratings for twelve, mainly 'intellectual,' objectives. In most cases higher evaluations were given when teachers had more able and mature pupils in mind.

CONCLUSIONS.

(1) In general, teachers preferred those objectives which might be described as 'intellectual' and 'subject-centred' in content to those which might be described as 'socio-emotional' and 'aesthetic'.

(2) Instances of agreement between the various sub-groups of teachers in their rating of the objectives far outnumbered those of disagreement. However—

(i) women teachers were consistently more generous in their ratings than men teachers and more inclined to favour child-centred objectives than were the men ; and

(ii) grammar school teachers were more inclined to rate highly intellectual and subject-centred objectives than were the other teachers.

(3) Evidence from a small number of teachers suggests that considerations of pupil's age and ability would have little effect upon the ratings of four-fifths of the objectives but quite an appreciable effect upon the remaining fifth which are mainly intellectual in content.

(*Manuscript received 19th August,* 1965)

THE EFFECTS OF ANXIETY ON LEARNING, TASK PERFORMANCE AND LEVEL OF ASPIRATION IN SECONDARY MODERN SCHOOL CHILDREN

By UNA MAGUIRE

(*Abstract of thesis submitted for degree of M.Ed. in the University of Manchester.*)

OBJECTS.

The main aim of this research was to discover the difference in the learning behaviour and task performance of secondary modern school boys and girls designated as 'High Anxious' (H.A.) and 'Low Anxious' (L.A.) on tests.

In addition, an attempt was made to obtain some information concerning the interaction effects of anxiety, introversion–extraversion and the various conditions under which tasks are performed.

An investigation was also made of the differential effects of success and failure on the levels of aspiration (i.e., goal setting) of the High Anxious and Low Anxious children.

METHOD

Three-hundred-and-thirty-five secondary modern school pupils, whose ages ranged from 13·0 to 14·10 (mean 13·10) were initially tested. The two girls' and three mixed schools where they attended were all situated in densely populated areas in or near Manchester, and the vast majority of the children came from working class homes.

The 335 children were given the following tests :

(i) The Manchester General Ability Test (Wiseman, 1959).

(ii) The Castaneda Questionnaire (Castaneda, *et al.*, 1956).

(iii) The Palmar Sweating Test (Silverman and Powell, 1944).

(iv) The Junior Maudsley Personality Inventory (Furneaux and Gibson, 1960).

From this large group of children small groups were chosen for the experimental work.

Two groups were defined as High Anxious (H.A.) and Low Anxious (L.A.), according to their scores on the Castaneda Questionnaire, and two further groups were so defined according to their scores on the Palmar Test.

Introverted and extraverted groups were defined according to their scores on the Junior Maudsley Personality Inventory.

Finally, four groups were chosen to take part in the experiment. The children in these groups were (1) H.A. introverts, (2) H A. extraverts, (3) L.A. introverts, (4) L.A. extraverts.

THE EXPERIMENTS.

Experiment 1 (*a*)—*Rote Learning.*

All the tasks used in the experiment were in Latin. Since none of the children had previously studied Latin, and since the range of I.Qs. was restricted (102–120), it was considered that any obtained differences in group scores would, if statistically significant, be due to the differential effects of anxiety or introversion–extraversion on the particular conditions under which the task was performed.

The first part of Experiment 1 (*a*) consisted of rote learning of Latin vocabulary (e.g., femina=woman) under different conditions as folows: ***neutral*** (without praise or criticism), ***ego-supported*** (with praise and encouragement), ***ego-undermined*** (with discouraging criticism).

The second part of Experiment 1 (*a*) consisted of rote learning of symbols under the same conditions (e.g., femina = Δ).

The children were later tested and the analysis of variance technique was used on the results to investigate main and interaction effects. (This same technique was used on all the test results).

Experiment 1 (*b*)—*Task Performance.*

In order to assess the children's performance on tasks other than rote learning, more difficult tests, which involved more complex learning, were devised. On this occasion the conditions were slightly different from those used for rote learning:

(i) *Task-Oriented*—every effort was made to concentrate the child on the task.

(ii) *Ego-Oriented, under slight stress*—every effort was made to involve the children in the task and there was ' time stress.'

(iii) *Ego-Oriented, under severe stress*—again every effort was made to involve the children, and in addition to ' time stress ' there were threats concerning what would happen should the child not do well enough.

RESULTS OF TESTS.

Experiment 1 (*a*)—*Anxiety as measured by Castaneda Questionnaire.*

(i) No clear pattern emerged from the results relating to the performance of the H.A. and L.A. groups. However, in one set of results, the boys' performance on the test of vocabulary learning, there was an interaction, statistically significant (at the ·05 level) of anxiety with introversion-extraversion. The H.A. introverts performed better than the L.A. introverts, and L.A. extraverts performed better than H.A. extraverts. This, however, was not so, in the other three sets of test results.

(ii) There was no significant difference between the performance of High Anxious and Low Anxious girls on this test. The ego-supported condition produced the best test results, and the ego-undermined condition the worst, the overall differences between the conditions being statistically significant at the ·01 level. Further analysis showed that the difference between the ego-supported and ego-undermined conditions was statistically significant at the ·01 level. The other two pairs of conditions did not differ significantly.

Experiment 1 (*a*)—*Anxiety as measured by Palmar Test.*

(i) No clear pattern emerged. In both vocabulary and symbol learning tests the H.A. girls performed better than the L.A. girls. In the vocabulary test the difference was statistically significant at the ·05 level, and in the symbol test the difference just failed to reach significance. In only one set of results, that of the girls, on the symbol test, were the overall differences between the conditions statistically significant (at the ·01 level). Again, the ego-supported conditions were shown to be superior.

Experiment 1 (*b*)—*Task Performance—Anxiety as measured by Castaneda Questionnaire.*

(i) In all four sets of results introverted children performed better than extraverted children.

(ii) In three of the four sets of results H.A. children performed better than L.A. children.

(iii) There was a tendency for H.A. introverted children to perform better than L.A. introverted children, and for L.A. extraverted children to perform better than H A. extraverted children.

(iv) Task-oriented conditions produced the best results, and the ego-oriented severe stress conditions the worst in three out of the four sets of results.

Anxiety as Measured by Palmar Test.

(i) There was no clear pattern of results. The boys' performance on Task Performance (difficult), however, showed that H.A. introverted boys performed better than L.A. introverted boys, and L.A. extraverted boys performed better than H.A. extraverted boys.

(ii) The task-oriented conditions produced the best test results and the ego-oriented severe stress the worst. The difference between the conditions was statistically significant at the ·05 level.

Experiment 2—Levels of Aspiration.

A series of fourteen tests was given to the children. Seven were given under conditions involving success and seven under conditions involving failure.

The failure tests consisted of short codes which the children were required to break in a very short period of time. After each test the children were asked to estimate their score on the following test: this was called the 'expected' score. From the difference between this score and the obtained score, discrepancy scores were calculated which were used in the interpretation of the test results.

The same procedure was used in the success conditions tests. These tests, which were quite simple, required the child to supply some missing letters in order to complete a word, or to give the meaning of a word.

RESULTS OF TESTS.

Experiment 2—Levels of Aspiration.

Realism in Levels of Aspiration—Conditions of Success.

(*a*) *Anxiety as Measured by Castaneda Questionnaire.*

   (i) L.A. children were on average, more realistic in their levels of aspiration than H.A. children, and extraverted children were more realistic in their levels of aspiration than introverted children, but the differences were not statistically significant at the ·05 level.

   (ii) Under conditions of success most children became less realistic.

(*b*) *Anxiety as Measured by Palmar Test.*

   (i) Here the H.A. boys were, on average, more realistic in their levels of aspiration than L.A. boys, though extraverted boys were more realistic in their levels of aspiration than introverted boys.

   (ii) The High Anxious extraverted girls were slightly more realistic in their levels of aspiration than the Low Anxious introverted girls, who, in turn, were more realistic than Low Anxious extraverted girls.

Realism in Levels of Aspiration—Conditions of Failure.

(*a*) *Anxiety as Measured by Castaneda Questionnaire.*
L.A. children were on average more realistic than H.A. children. As the tests progressed all children became less realistic in their levels of aspiration.

(*b*) *Anxiety as Measured by Palmar Test.*
- (i) H.A. groups were, on average, more realistic in their levels of aspiration than L.A. groups.
- (ii) Extraverted girls were more realistic in their levels of aspiration than introverted girls, the difference being statistically significant at the 5 per cent. level.

Rigidity in Levels of Aspiration—Conditions of Success.

(*a*) *Anxiety as Measured by Castaneda Questionnaire.*
There was a general tendency for L.A. children to be more rigid in their levels of aspiration than H.A. children. The tendency with the boys appeared to be for the introverted boys, and for girls with the extraverted girls.

(*b*) *Anxiety as Measured by Palmar Test.*
Here there was a general tendency for H.A. children to be more rigid in their levels of aspiration than L.A. children, but the difference was not statistically significant at the ·05 level.

Rigidity in Levels of Aspiration—Conditions of Failure.

(*a*) *Anxiety as Measured by Castaneda Questionnaire.*
L.A. boys were more rigid in their levels of aspiration than H.A. boys, but the difference was not statistically significant at the 5 per cent. level. The greater rigidity of the L.A. girls was confined to the introverted group.

(*b*) *Anxiety as Measured by Palmar Test.*
H.A children were more rigid in their levels of aspiration than L.A. children, and this was particularly so for the introverted groups.

(*Manuscript received 19th August, 1965*)

THE OPINIONS OF SCIENCE TEACHERS ON THE OBJECTIVES OF TEACHING SCIENCE

By GERARD J. POLLOCK

(*Abstract of thesis submitted for the degree of M.Ed. in the University of Manchester,* 1964)

The main aim of the research was to investigate the opinions of science teachers on the objectives of teaching science, and the dependence of these opinions on various parameters of the teaching population. Subsidiary aims were to ascertain the educational attitudes of science teachers, to correlate these attitudes with ratings of the science objectives, and to compare the opinions of science teachers with those of training college science lecturers.

Random samples of teachers and lecturers were drawn from the membership lists of the Science Masters' Association, the Association of Women Science Teachers, and the Association of Teachers in Colleges and Departments of Education. A 63 per cent. response was obtained.

Chi-square analysis was used to compare the respondents with the general populations of science teachers and training college lecturers. The sample of respondents consisted of seventy-six male and ninety-two female science teachers, twenty-six male and fifteen female training college science lecturers. The sample of science teachers could not be regarded as fully representative of the teaching population since it contained an excess of honours graduates concentrated in grammar school teachers and physics teachers. The training college lecturer sample could be regarded as a representative sample of the training college lecturer population.

Two questionnaires were used to obtain quantitative scores. The first was a rating sheet compiled by the writer containing sixty-five of the more common objectives of teaching science. The respondents rated each objective individually, on a five-point scale ranging from very desirable to undesirable. The responses were given scores ranging from 3 to –1. The sixty-five objectives could be arranged in five groups, ' Knowledge,' ' Interests,' ' Attitudes,' ' Skills ' and ' Cultural.' The second questionnaire was Oliver's " Survey of Opinions on Education," which measured the educational attitudes of the respondents on three dimensions, ' Naturalism-Idealism ' (N), ' Radicalism-Conservatism ' (R), and ' Tender-mindedness-Tough-mindedness ' (T).

The science objectives were generally rated high by female teachers and further education teachers, and low by physics teachers, university teachers and honours graduates. The teacher sample rated the objectives significantly higher than the training college lecturer sample. The various sub-groups of the samples showed very good agreement over the relative ratings of the objectives. The ' Attitudes ' and ' Skills ' objective groups were rated high, while ' Knowledge,' ' Interests ' and ' Cultural ' objective groups were rated low.

The agreement among the sub-groups of the samples was measured by Kendall's Co-efficient of Concordance (W) and Ebel's Intra Class Correlation Coefficient (E). The teacher sample gave values of $E=0{\cdot}947$ and $W=0{\cdot}851$. The training college lecturer sample gave values of $E=0{\cdot}885$ and $W=0{\cdot}797$. Very good agreement was found between both samples over the individual ratings of the objectives : $E=0{\cdot}960$, $W=0{\cdot}953$.

Significant interaction effects were found in the teacher sample among the ratings of the Science Subject groups, and the Range of Classes groups. No interaction effects were found among the training college lecturers The majority of the interaction effects concerned Biology teachers and Form I—IV teachers. Both groups considered it relatively important to arouse the child's interest in nature, to develop good leisure pursuits, and to secure health. Lower form teachers were also concerned to remove superstition and error from their pupils' minds, but were not greatly concerned with preparation for future courses in science. Both samples considered that producing a better understanding of our environment was the most important objective of science teaching. Objectives emphasising the importance of an open, unbiassed mind were rated highly by both samples, as were objectives emphasising accuracy and observation. Objectives emphasising vocational training, preparation for future courses in science, developing interest in scientific literature and scientists were rated lowest by the samples.

The two samples gave normal distribution of N, R, and T scores. Both had mean scores biassed towards idealism and radicalism, but differed on the T-scores where the teacher sample was biassed towards tough-mindedness, while the training college lecturer sample had a mean biassed slightly to tender-mindedness. The training college lecturers were significantly more tended-minded than the teacher sample. Teachers aged 50 and over were also more tough-minded than younger teachers.

Criterion groups based on the N, R, T scores were set up, consisting of the top and bottom fifty-seven respondents on each attitude dimension. The ratings of the science objectives by these criterion groups were analysed also by variance technique. The science objectives were rated higher by the ' idealist,' ' radical,' and ' tough-minded ' groups than by the ' naturalist,' ' conservative,' and ' tender-minded ' groups, respectively. The ratings by the N. and T. criterion groups, gave similar interaction effects. The ' idealist ' and ' tough-minded ' groups emphasised objectives dealing with a vocational impulse in teaching, accuracy, and truthfulness. The ' naturalist ' and ' tender-minded ' groups emphasised appreciation of the contributions of scientific method.

(*Manuscript received 19th August*, 1965)

THE TEACHING OF ELEMENTARY SCIENCE BY TELEVISION AND OTHER METHODS

By H. BARRINGTON.

(*Summary of thesis submitted for the degree of M.Ed., University of Manchester*, 1964)

AIMS.

The main aim was to compare the achievement of pupils who have studied physical science by television with that of others using conventional methods. A secondary concern was to obtain an objective measure of the attitude of pupils to being taught physical science by the medium of television.

THE SAMPLE.

An incidental sample was taken from the first and second forms of a four-form entry secondary modern school in the City of Manchester. The groups were similar in size (thirty) and contained nearly equal numbers of boys and girls. The standard deviations of ability of the groups were similar and initial differences in ability between groups were held constant in the statistical analysis.

THE METHOD

A.—Achievement.

An initial measure of the general ability of the pupils was obtained by combining scores on a Moray House Verbal Reasoning test and Raven's Matrices (1938) test. (Called the X score.) A measure of the pupils' initial knowledge of the subject matter of the science lessons to be taken was obtained from an objective test of 100 items designed to cover the content of the course to be followed (Test No. 1).

Four groups of pupils then studied the same subject matter by different methods for five weeks. One group was taught by using television lessons (from the B.B.C. series " Discovering Science ") integrated with laboratory lessons. A second group had conventional laboratory lessons, a third was taught by taking television lessons (repeat of the above B.B.C. lessons) in the classroom and a fourth had conventional classroom lessons based on a textbook. A fifth group was used as a control group to detect any incidental learning which might take place through other educational influences such as general television programmes viewed at home. This group followed a series of biology lessons unrelated to the physical science lessons taken by the other groups.

It was found administratively impossible to employ different teachers or to interchange teachers in order to nullify the teacher variable. The writer consequently arranged to take all groups and administer all tests personally. After the teaching period the 100 items test was completed by all pupils for a second time (Test No. 2). The mean change in score of each group was called the Y score.

A product moment correlation coefficient indicated a definite relationship (·4) between initial ability and achievement after the teaching period, i.e., between X and Y. Analysis of covariance was, therefore, employed using X scores as the independent or predicting variable and Y scores as the dependent variable The F test, on the variances obtained from the adjusted sums of squares on the Y scores, proved highly significant ($P < \cdot 01$). *t* tests were, therefore, carried out on the adjusted Y means.

From each of the four teaching groups it was found possible to select seven pupils of approximately similar ability to pupils in the other groups, *t* tests were carried out on the mean changes in scores of these matched groups and comparisons were drawn with the results of the covariance analysis.

The effect of a teaching method may be more accurately assessed if both the amount of learning achieved by the end of the teaching period and that retained after a longer interval of time are measured. Consequently, four months after the end of the experimental teaching period, the 100 items test was taken by all pupils for a third time (Test No. 3).

t tests were then carried out to determine whether differences between retention scores (Test 2—Test 3) were significant.

B.—Attitude.

After try-out of fifty items, in the form of a summated ratings questionnaire, on pupils in three secondary schools which had used the " Discovering Science " television lessons, analysis of responses produced twenty-four items for a Likert-type attitude scale. The questionnaire was assumed to be valid by definition and its reliability (split half with Spearman-Brown correction) was ·81.

The questionnaire was completed by the two groups which had used television during the experimental teaching period.

FINDINGS.

(1) The scores of the control group suggested that learning of the subject matter was not facilitated by external agencies.

(2) No significant difference was found between the achievement of the T.V. / laboratory group and the conventional laboratory group.

(3) The achievement of the T.V. /classroom group was significantly higher (·05) than that of the conventional classroom group.

(4) No actual difference was found between the loss of learning of the T.V. / laboratory group and that of the conventional laboratory group.

(5) No significant difference was found between the loss of learning of the T.V. / classroom group and that of the conventional classroom group. Both groups suffered a similar loss of learning—about 30 per cent.

(6) Each television group reported a favourable attitude to learning science by television. No significant difference was found between the attitudes of the two groups.

(7) Other findings were :

(*a*) Some teaching is better than none at all.

(*b*) Television science lessons can provide the means for a quick gain in information.

(*c*) Television science lessons require adequate follow-up lessons if learning is to be retained.

(*d*) The attitude of pupils to television science lessons is, in general, favourable but depends partly on the pupils' ability level The less able pupils are enthusiastic but the more able the pupils, the less favourable their attitude.

(*e*) The pupil's attitude to being taught by television science lessons does not appear to affect his performance on an achievement test.

(*f*) Television lessons hold the pupil's attention.

(*g*) Television science lessons save valuable time for the teacher.

The findings of this research must be considered as indicative rather than conclusive and generalisations must be restricted and tentative. It should be noted, however, that supporting evidence is provided by a large number of instructional television researches carried out in the United States of America.

(*Manuscript received* 22*nd November*, 1965)

AGE AND MENTAL ABILITY: RE-TESTING WITH THIRTY-THREE YEARS' INTERVAL

BY ROBERT B. BURNS

(*Abstract of Thesis submitted in part fulfilment of requirements for the degree of Bachelor of Education in the University of Aberdeen*)

AIM.

To extend a previous longitudinal study of the relationship between age and mental ability (Nisbet, J. D., *Brit. J. Educ. Psychol.*, 27, 190-8, 1957).

SUBJECTS.

Eighty university graduates (thirty-six men and forty-four women, mean age 56 years) who had already been tested on two previous occasions—during teacher training in Aberdeen in the period 1930-33, when their mean age was 22½ years; and in 1955, when their mean age was 47 years. Eighteen other graduates who had taken the original test in 1930-33, but had not taken the 1955 re-test, were used as a small control group to assess test sophistication.

TEST.

A shortened version of the Simplex Group Intelligence Test (from sub-test 3B to the end), containing two vocabulary sub-tests, seven verbal reasoning tests, three tests involving number (number series, digit memory and number relations), a mirror images test and a substitution or coding test. In 1955 and in 1964 the subjects did the test unsupervised in their own homes according to special directions.

RESULTS.

The mean score of the eighty graduates in 1964 was 74·8, compared with 77·8 in 1955, and 67·9 in 1930-33. The improvement in mean score between 1930-33 and 1955 had proved significant and was equivalent to about 0·7 of the standard deviation of scores in an unselected adult population. Between 1930-33 and 1964 the improvement was again significant and equivalent to 0·44 σ. Between 1955 and 1964, there was a significant decline in mean score equivalent to 0·18 σ. Performance in verbal and vocabulary sub-tests continued to show improvement in this later period, that is, between the ages of 47 and 56. But mean score in the number sub-tests, having improved up to age 47 (though less markedly than the verbal scores) showed a sharp decline in the period from age 47 to 56. Results from the small control group indicated that practice on this test had not affected results. Analysis of scores of a small group of mathematics teachers showed that their performance in number items had been better sustained, while language teachers obtained higher scores than others on verbal items.

CONCLUSIONS.

Some at least of the changes in scores were attributable to genuine changes in ability due to continued learning in the intellectually stimulating environment of teaching, the occupation still followed by sixty of the eighty graduates. The decrease in mean score between age 47 and age 56 was due largely to poorer performance on number items. Previous longitudinal studies covering a shorter age span had shown that performance improved less with age in number tests than in verbal tests. In this group, in their fifties, performance on number had deteriorated. The relative stability of scores in verbal and vocabulary sub-tests may be due to continuing practice and additional learning in these skills, preventing or concealing deterioration.

(*Manuscript received 7th July*, 1965)

THE RELATIONSHIP BETWEEN READING ABILITY AND VERBAL INTELLIGENCE TEST PERFORMANCE

By CONAL STUART RILEY

(Abstract of a thesis submitted as part of the requirement for the award of Diploma in Applied Psychology from the University of Liverpool.)

AIM.

To investigate the relationship between performance on a written verbal intelligence test and performance on an individually administered intelligence test which included both verbal and performance tasks, and to examine the influence of reading ability on these test performances.

THE CHILDREN STUDIED.

The investigation was carried out with eighty-three children in the 8–9 age group who comprised the second year of the junior school section of a Liverpool primary school.

METHOD.

Three tests were used: the Graded Reading Vocabulary Test R.1[1], the Essential Intelligence Test[2], and the Wechsler Intelligence Scale for Children[3]. The reading test and the written intelligence test were given to as many children as were available at the time of testing, N=83 and 78, respectively. The Wechsler Intelligence Scale was given to a sample (N=38) of the children which was drawn at random and proved to be representative of the group as a whole.

RESULTS.

The group was divided into 'above the median' and 'below the median' readers (Median R.Q.=93) and for each group the correlation between Essential Intelligence Test and W.I.S.C was calculated.

For 'above the median' readers the two tests correlated +0·89 (P=0·01). For 'below the median' readers, however, the correlation was 0·38 (N.S.).

Further, a Chi-square test was carried out to ascertain the relation between the I.Q. values given by these two tests. The test showed ($X^2=7{\cdot}2$ P=0·01) that discrepancies were in opposite directions for the two groups. For the 'above the median' group Essential Intelligence Test I.Q. exceeded W.I.S.C. I.Q. while for the 'below the median' group the relationship was reversed.

It was also found that the scatter diagram showing the relationship between Reading Age and Essential Intelligence Test I.Q. did not show any marked discontinuity at Reading Age=7½ years. This suggests that the statement by Schonell and Adams[2] in the test manual that, "the test should not be given to any children with Reading Ages less than 7½ years" is not a necessary condition of test use.

CONCLUSIONS.

If the W.I.S.C. I.Q. is accepted as criterion, the Essential Intelligence Test I.Q. is a good predictor only for 'above the median' readers.

The Essential Intelligence Test tends to overestimate the scores of 'above the median' readers. The fact that in this study the Essential Intelligence Test also underestimated the scores of 'below the median' readers may have been due to the fact that the Reading Age tables had to be extrapolated beyond the lowest level given in the manual to take into account some low scoring subjects.

REFERENCES.

[1] SCHONELL, F. J., and SCHONELL, F. E. (1962). *Diagnostic and Attainment Testing.* London: Oliver and Boyd.

[2] SCHONELL, F. J, and ADAMS, R. H. (n.d.). *The Essential Junior Intelligence Test* (Manual of Directions). London. Oliver and Boyd.

[3] WECHSLER, D. (1949). *Wechsler Intelligence Scale for Children* (Manual). New York: The Psychological Corporation.

(Manuscript received 9th September, 1965)

THE ASSOCIATION OF AUDITORY HIGH FREQUENCY WEAKNESS WITH VERBAL AND WRITTEN COMPREHENSION AND EXPRESSION

BY G. F. REED

(*Abstract of thesis presented for the degree of M.Litt. of the University of Durham*, 1964)

INTRODUCTION.

Clinical observation prompted the hypothesis that children who fail to respond to high frequency tones outside the speech range (and who may thus be unable to discriminate some of the components of consonantal sounds) may show retardation or distortion in the development of their understanding and articulation of speech. From this followed the predictions that the general language development of such children would be hindered as would their acquisition of reading and of skill in written expression.

PROCEDURE AND RESULTS.

Audiometric study of groups of children to ascertain their upper limits of hearing was carried out, using a Marconi Audio Tester coupled to a high fidelity speaker and tweeter unit. This apparatus allowed the reproduction of pure tones of up to about 22,000 c.p.s. at a maximum intensity of 50 db.

A.—The Main Survey.

Using area/stratified sampling, a group of 209 children (119 boys and 90 girls) was selected as being representative of their age range (7 to 15) in a large city (total school population 46,000). Audiometric survey of this group was undertaken in order to establish provisional norms of response to the test equipment. The survey produced data which may have practical and theoretical importance in their own right:

(*a*) The distribution of responses in each age group was continuous and high inter-subject consistency was recorded for all groups.

(*b*) No significant differences were found between sexes or between geographic areas (which had been selected as being associated with socio-economic conditions).

(*c*) A significant lowering of upper frequency response limits was found to occur with age from as early as twelve years.

(d) With the equipment in use it was found that some children responded to tones of up to 21,000 c.p.s., but the mean from age 7 to age 12 was just above 18 kilo-cycles, dropping to just above 16 kilo-cycles at age 15.

B.—Three Subsidiary Surveys.

The results of three further surveys of special groups of children were in general agreement with the above findings. These surveys were carried out to investigate whether auditory high frequency weakness (HFW) is associated with intelligence and personality variables, and whether it might be a factor in severe reading retardation.

(*a*) Among eighty-six children currently attending a child guidance clinic, a positive and statistically significant association was found between high frequency auditory acuity and intelligence as determined on an individual mixed intelligence test (the Terman-Merrill Revision of the Binet).

(*b*) A significant difference was found between the HF auditory acuity of children who had been psychiatrically classified as ' Anxiety State ' and those classified as ' Hysteric ' The operational criteria used in these diagnoses makes it permissible to regard the two clinical categories as being equivalent to ' Introverted' and ' Extraverted.' When the effects of intelligence were discounted, the superiority of the first group still held.

(*c*) A group of thirty-five severely retarded readers demonstrated auditory acuities of wider scatter than the normal group, twenty-one having ' Upper Auditory Limits ' below the means of their age groups on the main survey. But this was found to be associated with the distribution of personality types in the group. Those retarded readers who were classifiable as ' Anxiety State ' or introverted achieved high U.A.L.s. This suggests that HFW is more common among severely retarded readers than among children of normal achievement, but that this may be a function of the distibution of ' Hysteric ' or extraverted children among such groups.

(*d*) A group of forty educationally subnormal children contained a large proportion of cases of relative HFW which is in line with findings (*a*) and (*b*) above. Many of these children would be classifiable as ' Hysteric,' whilst their ascertainment as E.S.N. is based upon their low intelligence. Those members of the group with I.Qs. below 65 were generally unable to respond to high frequency testing with reliability (because of lack of understanding of the task or hyper-distractibility).

C.—*The Experimental Study.*

From a total of 464 children given audiometric tests (twenty children regarded as being cases of HFW were identified, using as a criterion a depression of U.A.L. response of more than two standard deviations below the appropriate age ' norms ' found in the main survey. A ' control ' was found for each experimental child, matched in terms of age, sex, socio-economic status, non-verbal intelligence and school class. The ages of these children from 8 to 13, and their intelligence levels, as gauged by the Progressive Matrices, ranged from the 75th percentile down to below the 5th percentile. The speech, language (verbal and written) and reading abilities of the children were then assessed, using established tests where such were available. It was thus possible to compare the performance of the experimental cases with that of their matched controls and also, in most instances, with standardized norms. The results showed that in many tests both groups were of below average ability. But the children in the experimental group proved to be inferior to their controls in seven out of eight tests.

(*a*) The most significant difference between the groups was in the results of the test of speech articulation (the Francis Phonetic Word List). Whereas none of the control group made any errors whatsoever, thirteen of the experimental group proved to have some degree of articulatory defect, two being very serious. Two of those HFW children who made no error on this articulation test showed faulty intonation, and another two were amongst those who had received speech therapy in the past. So that a total of seventeen of the twenty HFW cases had some current defect of speech or had displayed defect in the past.

(*b*) On both the Mill Hill Vocabulary Scale and the Watts' English Language Scale no member of the HFW group achieved a better score than did his matched control, the differences between the groups being highly statistically significant on both tests.

On the Vocabulary test ten of the HFW group were below the 5th percentile for their age, and only three of the group scored at or above the 50th percentile. On the Language test fourteen of the HFW group achieved Language Ages of less than 8, whereas sixteen of the controls achieved a L.A. of 8 or above.

As a group, the HFW subjects showed pronounced immaturity in both the acquisition of vocabulary and in oral language.

(*c*) Free compositions written by the subjects were analyzed, using as measures the numbers of words written and the percentage subordinate clauses (La Brant). Applicable standard for these two measures have not been published, so similar compositions were drawn from small samples of local school children at each relevant age level for comparative purposes.

(i) The HFW children wrote far fewer words than either their controls (in fifteen cases) or the sample means. Differences were significant in both cases. Six HFW children wrote a number of words which was three or more standard deviations below the appropriate sample means.

(ii) When clauses were analysed, five of the HFW group's compositions had to be excluded because they were so primitive as to be incomprehensible. In only one HFW case was the subordination index superior to that of his control, and in only two cases was it above the sample mean. Differences between means were highly significant.

The HFW children as a group showed not only a paucity of output but a paucity of ideas, and clumsiness and immaturity of language usage. Their short, disconnected phrases were flat and uninteresting by comparison even with those of their controls.

(*d*) On the Schonell Test of Sentence Structure eight of the HFW group made no score at all, and only two made a better score than their controls. The group difference was highly significant. Again, the HFW group proved to be grossly retarded in grammatical usage as expressed in their ability to handle conjunctions and inversions and to substitute for simple sentences adjectival or adverbial phrases or clauses.

(*e*) Schonell's word recognition test was the only one in this battery where the HFW group's results, although inferior to those of the control group, were not significantly so Only one HFW child was totally unable to read, but nine others had Reading Ages two or more years below their chronological ages. However, four of them achieved R.As. beyond their C.As.

On the Kelvin test of reading *comprehension* the control group was again superior• the difference being highly significant. Only three HFW children achieved results above the 50th percentile, whilst six were below the 5th percentile. This highlights differences between the groups in verbal reasoning and the comprehension of language.

The lack of significant difference between the groups on the test of mechanical reading may be due to the fact that no less than ten of the HFW group had received remedial help with their reading. Regular testing with word recognition tests during such courses of specialist tuition has been shown to impart facility with this particular type of test which does not spread to other, more crucial forms of reading ability.

(*f*) The HFW group had also received many types of educational, medical or psychological help or attention, such as special classes or groups, speech therapy, E.N.T. examinations, child guidance, clinic or school psychological service referrals or E.S.N. examinations. Only one HFW child had not been referred in some way, whilst fifteen of the twenty had been referred for more than one type of help. As a group the HFW children had drawn attention to themselves throughout their school lives as being retarded or abnormal in speech, hearing, language, reading or behaviour, or, more commonly, in some combination of these.

DISCUSSION.

The association found between HF ability and intelligence and personality type, makes it unlikely that relative HFW is a simple sensory loss. It is suggested that insufficient attention has been paid to pure tone audiometry as a psycho-physical procedure. Responses to pure tone stimuli are not necessarily a crucial index of 'hearing' ability. They reflect various psychological variables. A fruitful approach to the present problem, may, therefore, be from the viewpoint of experimental psychology, and it is suggested that studies of 'vigilance' are of direct relevance. Performance in perceptual motor tasks where responses are required to faint and intermittent stimuli are known to be associated with extraversion/introversion and the present findings are, therefore, discussed in terms of 'set,' response consistency and reactive inhibition. HFW may thus be regarded not as a sensory weakness causing retardation of linguistic development, but as having the status of an intervening variable or concomitant.

However this may be, the results of this study have given support to the original hypothesis and substantiated the predictions drawn from it. Evidence has been presented to show pronounced association between auditory high frequency weakness and retardation or distortion in the understanding and articulation of speech, and immature language development, including gross retardation in reading skills and written expression. Although relatively rare the condition may be seen to be of pronounced clinical significance and a potentially important source of educational backwardness.

(*Manuscript received 2nd June*, 1965)

A COMPARISON OF THE ERRORS MADE BY GRAMMAR AND TECHNICAL SCHOOL PUPILS IN ALTERNATIVE PAPERS IN G.C.E. ENGLISH LANGUAGE

By D. M. BRANNAN

(*Summary of thesis submitted for the degree of M.Ed., University of Manchester*, 1963)

AIMS OF THE INVESTIGATION.

The aims of the investigation were: (*a*) to identify the mistakes made in the essays written by two groups of candidates who were entered for alternative English Language papers in the Northern Universities' Joint Matriculation Board General Certificate of Education examination; (*b*) to assign a score to these mistakes, according to their importance; and (*c*) to discover whether the greater knowledge of formal grammar of one of the two groups of candidates was reflected in the technical accuracy of their written work.

METHODS OF INVESTIGATION.

The pupils were seventy-four boys from a secondary technical high school and eighty-seven boys from a secondary grammar school. The two groups, of the same average age, were in the final year of a five-years course leading to the General Certificate of Education examination. The technical school group were the whole of the school's fifth year. The grammar school boys constituted the middle and lower streams of their year: the top class took the General Certificate of Education examination after a four years' course and were, therefore, significantly younger than the boys from the other fifth forms.

In the examinations taken at the end of the school year, 53·8 per cent. of the grammar school boys were successful in Paper 'A' of the General Certificate of Education English Language examination; 52·9 per cent. of the technical school candidates passed in Paper 'B' of the same examination. In so far as the General Certificate of Education English Language examination adequately tests general English ability, and in so far as the two English Language papers are of equal standard, the pupils of the two schools appear to be of similar attainment in English.

The two groups were tested for intelligence by means of a group test of general intelligence (AH4). No significant difference was found.

The technical school course included little English grammar—in the first two years an elementary consideration of the parts of speech and the structure of the simple sentence, in the remaining years incidental references to verb, adjective, pronoun and conjunction. The grammar school course included a continuous and much more detailed study of the parts of speech and the analysis of complex and compound sentences. The two groups were tested for knowledge of English grammar by means of an unpublished test of grammar constructed by Robinson (1959). The scores obtained by the grammar school pupils were significantly higher than those obtained by the technical school pupils at the 0·01 level.

Each boy was expected to produce three essays, two as part of the normal school course and a third in the General Certificate of Education English Language examination. The two English Language papers provided the titles for one of the essays; the titles of the other two essays were selected from a common list of twelve. In all, 433 essays were written (196 by the technical school group and 237 by the grammar school group; not all the boys produced three essays).

The essays were examined and a count was taken of the technical errors they contained. Thirty-nine kinds of mistake in grammar and punctuation were identified. The list of errors, with examples, was given to twenty-five experienced teachers of English who were invited to indicate the importance they attached to each kind of error. The judgment was expressed in terms of a six-point scale—' O ' indicating that the mistake was considered to be of no significance at all and ' 5' that the mistakes were considered to reveal a serious weakness in ability to write acceptable English. Some two months after completing their assessments the twenty-five teachers were asked to repeat their judgments. Product moment correlation coefficients for the two sets of assessments varied from 0·437 to 0·936. The judgments of each of twelve teachers produced reliability corefficients of 0·7 and above. These twelve pairs of assessments were combined and used to weight the thirty-nine errors for importance. The mistakes of each pupil were weighted and added to produce an error score for each essay and a total score which was adopted as the measure of his ability to write correct English. The component scores for grammar errors and for punctuation errors (in each essay) were also calculated.

RESULTS.

The mistakes identified in the written work of the two groups, with their frequency and weight, are given below..

	Frequency		
	Grammar School	Tech. School	Weight
SPELLING MISTAKES	800	886	
MISTAKES IN STRUCTURE:			
Omission of a fullstop between separate sentences	554	377	4·33
Gross mistakes in the use of the comma	189	112	2·87
Intrusion of a full stop between parts of the same sentence	78	52	4·29
Absence from a sentence of a finite verb	74	45	3·71
Badly related phrase or clause	43	40	3·46
Unrelated dependent clause	5	20	4·04
Mistakes in the use of the semi-colon	19	11	3·21
Absence from a sentence of a subject	6	4	3·87
MISTAKES IN THE USE OF VERBS:			
Lack of agreement in number between subject and verb	64	64	3·33
Unjustified change of tense	30	46	2·87
Miscellaneous mistakes in the use of verbs	33	21	3·37
Mistakes in the use of ' would ' and ' should '	21	32	0·96
Mistakes in the use of participles	15	26	3·91
Wrong sequence of tenses	9	26	2·71
Mistakes in the use of ' shall ' and ' will '	18	11	1·37
Unrelated or badly related participles	11	16	3·04
The fused participle	5	6	2·41
Mistakes in mood	2	4	1·54

MISTAKES IN THE USE OF PRONOUNS :			
Lack of agreement in number between a pronoun or a possessive adjective and its related noun	83	69	2·96
Vague use of pronouns	48	47	3·12
Mistakes in the use of ' who ' and ' which '	34	29	3·96
Inconsistency in the use of ' one '	9	7	2·83
Mistakes in the case of pronouns	11	4	3·62
Mistakes in the case of relative pronouns	6	5	3·37
MISTAKES IN THE USE OF ADJECTIVES AND ADVERBS :			
Confusion of adverb and adjective	23	20	3·33
Mistakes in the use of the comparative and superlative	3	4	3·82
Use of ' that ' as an adverb of degree	2	2	2·83
MISTAKES IN THE USE OF PREPOSITIONS AND CONJUNCTIONS :			
Mistakes in the use of the preposition	68	68	3·04
Use of ' like ' as a conjunction	5	7	2·25
Mistakes in the use of correlative conjunctions	3	7	2·96
Mistakes in the use of ' due to '	4	4	1·62
Mistakes in the use of the construction ' the reason is that. . .'	3	4	2·1
MISTAKES IN PUNCTUATION :			
Mistakes in the use of the hyphen	207	165	2·04
Mistakes in the use of the capital	232	130	3·04
Mistakes in the use of the question mark	41	31	2·87
Mistakes in the use of quotation marks	27	14	2·96
MISTAKES IN THE USE OF THE APOSTROPHE :			
Mistakes in the use of the possessive apostrophe	218	213	3·17
Confusion of ' it's ' and ' its '	52	29	3·17
Misuse of the apostrophe with the elision of words	20	18	2·91

There were no significant differences between the means of error scores achieved by the two groups of candidates.

The product moment correlation coefficients for the error scores produced by the same children in different essays were as follows :

	No.	Grammar Scores	Grammar + Punctuation Scores
G.C.E. Essay and Second Essay	150	·6797	·5466
G.C.E. Essay and Third Essay	120	·7081	·5633
Second Essay and Third Essay	122	·7003	·5623

CONCLUSIONS.

(1) Errors in spelling, punctuation (including the use of the apostrophe) and sentence structure are the most frequent types of error. These categories of error accounted for 82 per cent. of all mistakes.

(2) Certain errors which occur infrequently are listed high in a rank order of importance according to the assessments of twelve judges. These mistakes are considered to represent a more serious deficiency in powers of expression than their frequency suggests.

(3) The difference between the two groups in knowledge of formal grammar was not reflected by significant differences in the means of error scores calculated for the two groups. In so far as these error scores adequately measure the ' correctness ' of written expression there appears to be little difference between the two groups in ability to write technically accurate English.

(4) There is some evidence in this investigation to support the notion of a ' vocabulary of error ' which is fairly constant in spite of differences in essay title.

(*Manuscript received* 1*st April,* 1965)

CRITICAL NOTICE

WALLACE, J. G. (1965). *Concept Growth and the Education of the Child.* London . National Foundation for Educational Research, pp. 268, 42.

BY E. A. LUNZER

The sub-title of this work is *A Survey of Research on Conceptualisation*, and the aim of its author is to provide " a necessary background to the highlighting of problems, subjects and methods which could usefully be incorporated in an experimental programme dealing with conceptualisation as an aspect of intellectual development in children." The book is divided into two sections. The first, consisting of only two chapters, is a review of recent studies " in the behaviourist mould " ; the second, and by far the larger, is concerned with what is called " the main stream of ontogenetic studies of conceptualisation " · the greater portion of it is taken up with a consideration of studies directly inspired by the work of Piaget, although it includes mention or summaries or discussions of many articles which owe nothing to Piaget, but bear on a similar content, e.g , the development of children's notions of time. In order to avoid a reduplication of the work of Vinacke and Russell Wallace has concentrated on work carried out since 1955. Nevertheless, the magnitude of his undertaking may be gauged by the fact that the bibliography extends to tweny pages and numbers more than 500 references, the majority of them being comparatively recent.

Wallace's work was undertaken as a contribution to the preliminary phase of a research project, and its publication is directed primarily to the research worker in the field. It is important to insist on this point in view of several criticisms which can be made and must be made. For there is no doubt at all in the mind of this reviewer that, to the reader for whom it is intended, this book must prove invaluable. The coverage is not perfect, but this is hardly surprising. Among the more serious omissions one notes an absence of reference to the work of Bartlett and the related studies of Donaldson and others, as well as the failure to mention the recent work of Witkin and Kagan on conceptual style or the related work on creativity which has been recently reviewed by Vernon in *Educational Research.* The first section is far more cursory in treatment than the second, and one notes for instance the failure to discuss the full implications of Skinner's notions of linear programming to conceptual learning and their application by Gagne and others. Indeed, one has the impression that quite sizeable portions of these two chapters may have been taken with little change in wording from review articles, notably Kendler's review in the *Annual Review of Psychology*, 1961. Again, in the second section, some areas receive far more complete treatment than others, and the chapter on conceptualisation in the handicapped barely scratches the surface of the vast body of theory and experiment that exists in this field. Nevertheless, it remains that Wallace has surveyed a very impressive array of literature (his chapter on Educational Aspects should be especially helpful), covering most of the topics that are relevant to his title, and he has arranged the material in a way which makes the book a most useful source of reference.

While the work should prove an admirable aid to the research worker, its value as an introduction to the topic is much more questionable. The major portion of the book is concerned with studies based on the work of Piaget or related to it. Yet the present reviewer can find no *evidence* that the writer has read a single major work of Piaget (although he may have done), although there is abundant evidence that he has made a thorough study of the growing corpus of derivative work. Thus, the chapter which introduces Piaget's work is very largely composed of extracts from a recent article by Inhelder (not always acknowledged), while the introductory sections dealing with specific topics like number, area, volume, classification, and so on, rely far too heavily on the relatively short summaries found in replication studies by Lovell and others. This imbalance in treatment is reflected in some of Wallace's appraisals. Thus, after devoting less than a page to a summary of Piaget's work on number, Wallace describes a number of less seminal studies in great detail. These include an early article by Elkind (1961), in which this author offers what purports to be a re-statement of Piaget's views, although in fact, the description is quite out of keeping not only with the letter of Piaget's work but with its whole spirit. Wallace's summary of this extends over nearly a page and a half, and the erroneous interpretations are reproduced without comment—where Flavell (1963, p. 384), reporting the same work, is careful to add in a footnote that both in their definition of terms and in their operational expression in the form of test questions there is a ' lack of concordance ' between Elkind and Piaget. Similarly, Wallace is enthusiastic in reporting the ' more original and thought provoking ' experiment of Braine (1959) in which this author tries to prove that children of 4–5 have an operational understanding of length Yet this study by Braine contains quite radical methodological weaknesses which completely invalidate its results, and these have been thoroughly exposed by Smedslund (1963). It should be added in fairness that in general, Wallace is more often content to report studies and conclusions, with comparatively little evaluation, and, so far as the present reviewer can judge, his reporting is nearly always accurate.

The last chapter is largely devoted to an evaluation of the work of Piaget and its potential for future research work. Much of it is devoted to reviewing the criticisms and appraisals contained elsewhere, including the important monograph produced by Kessen and Kuhlmann (1962) and the critical notices of *The Growth of Logical Thinking*, by Bruner and Parsons. Wallace himself makes an earnest and justifiable plea for more longitudinal studies. He also argues for the adoption of more objective methods of experimentation. While this latter demand is not without foundation, the reader is unlikely to find much help from Wallace's solution, which is to discover "a qualitative statistical method which will detect those categories or patterns determined by the data themselves without the influence of the investigator's preconceptions " (p. 217). It is sufficient to point out that scientific advance is always dependent on the choice of what controls should be applied and what effects are to be observed, and these in turn are invariably determined by the investigator's preconceptions, i.e., by the model of nature which he constructs and desires to verify. Thus, to take a case in point, Piaget's observations on the development of conservation depend on the construction of situations which will reveal that development, and such situations could not have been constructed without the appropriate model of the child's cognitive processes. It is perfectly true that future research should aim at more and more unambiguous ways of delineating the precise features of this and other forms of development, but this cannot be done by relying on some omnipotent ' qualitative statistical method ' to produce valid categorisations out of random data.

Finally, a word about the relation between the work of Piaget and the behaviourist approach to learning. As Wallace correctly points out, there are fundamental differences between the two standpoints, and these are bound up with the stress laid by Piaget on the activity of the subject in organising his experience. His psychology is, therefore, one that seeks to interpret behaviour in terms of the organisation which the subject imposes both on his activity and on his experience. This means that a stimulus is never a *mere* stimulus, since the readiness to modify action in accordance with this or that aspect of it is entirely a function of the subject's ongoing activity—or what Miller, Galanter and Pribram (1961) refer to as the Plan. Moreover, even the perception of the stimulus depends on the plan, since what the subject perceives depends on where he looks and how he looks, i.e., on the scanning which is an integral component of every plan—and this notwithstanding the very real sense in which not only sensation, but perception too, reflects a one-one correspondence with the pattern of physical stimulation (if it did not do so, it would be of little value, however systematic the plan). Similarly, the response is never a *mere* response, for it always represents some switching of behaviour in accordance with some plan—even where the plan in question is a higher order one which indicates that on receipt of the given stimulus the subject must abandon the ongoing lower level plan and initiate another (a psychologist writing a sentence in a book on conceptualisation would shift to an alternative plan if he chanced to see billows of smoke seeping in under the study-room door). Now if one adopts such an approach, it follows that by attempting to reach objectivity through identifying the behavioural stimulus with the physical stimulus, accepting the latter as given, and ascribing the response to a mere stimulus-response connexion, classical behaviourism manages to sidestep the central problem of psychology.

It is true that neo-behaviourism is more sophisticated (and neo-behaviourism is as old as Hull), for it leans heavily on the concept of the mediating stimulus. But this only shifts the problem of the origin of the stimulus into a corner, while retaining all the artificiality of the superstructure to which it gives rise—a superstructure which continues to be an encumbrance to its adherents (even though good insights and first-class experimentation are far from unknown in their ranks, as witness the work of the Kendlers, which Wallace quite rightly esteems highly).

Two things follow. The first is that attempts to render Piaget more palatable by translating him into a behaviourist jargon, notably by Apostel, Berlyne and Stevenson, are misguided. For, if it is even half true that Piaget's insights have succeeded in ushering in a new epoch in the study of cognitive development, then this is precisely because he is not hampered by the pseudo-objectiveism of S–R dogma. The second is that by a proper analysis both of Piaget's standpoint and of behaviourism one can avoid the error of lumping together all kinds of psychologists as 'in the behaviourist mould,' simply on the grounds that their experimentation is more rigorous than that of Piaget and they happen to be American. Such a description has little application to Bruner, and still less to Simon. In particular, such an analysis leads to the recognition that, even though there is little common starting ground between Piaget and behaviourism, there is a great deal that he shares with Bartlett and his followers in this country, and with cybernetically oriented psychologists both here and in the U.S.A. Indeed, the translation of Piaget's ideas into cybernetic terms is perfectly feasible, and the difficulties encountered in making it would be precisely such as would lead to the clarification of his model and the experimental testing of its details.

Much of the foregoing has been critical. It, therefore, behoves the reviewer to re-iterate the point that Wallace's book represents a very great deal of well-guided effort in combing the literature, and that it will certainly constitute a most profitable source of reference to workers in the field.

REFERENCES.

BRAINE, M. D S. (1959) The ontogeny of certain logical operations : Piaget's formulations examined by non-verbal methods. *Psychol. Monogr.*, 73, No. 5 (whole number 475).

ELKIND, D. (1961). The development of quantitative thinking : a systematic replication of Piaget's studies. *J. Genet. Psychol.*, 98, 37-46

FLAVELL, J. H. (1963). *The Developmental Psychology of Jean Piaget.* Princeton, N.J. and London · Van Nostrand.

KESSEN, W., and KUHLMAN, C. (Ed.). (1962). Thought in the Young Child. *Mon. Soc Res. Child Dev.*, 27, No. 2 (Serial No. 83)

MILLER, G. A., GALANTER, E. H., and PRIBRAM, K. H. (1960). *Plans and the Structure of Behaviour.* New York : Henry Holt.

SMEDSLUND, J. (1963). Development of concrete transitivity of length in children. *Child. Dev*, 34, 389-405.

BOOK REVIEWS

Berlyne, D. E. (1965). *Structure and Direction in Thinking.* New York, London, Sydney : John Wiley and Sons, Inc., pp. vii+378, 68s.

Berlyne's long theoretical essay is an important development in the psychology of thinking. His allegiance is to S–R associationism. His main analytical tools are the mediating response and the habit-family hierarchy. The brief treatment of productive thinking remains strikingly faithful to Hullian orthodoxy. But Berlyne has concentrated on *reproductive* thinking, i.e., thinking which " simply applies established knowledge and sets established routines in motion." His major theme is a careful exposition, in S–R terms, of the continuity between simple learned behaviour and the concrete and formal operations studied by Piaget. This entails the recognition of distinctions to which, heretofore, S–R theory has remained blithely insensitive.

The first innovation is a division of symbolic mediating responses (thoughts) into two types : situational and transformational. Thinking of the undirected, free-associational kind is regarded as an indeterminate sequence of situational thoughts. Directed, rational thinking, however, is seen as an alternating sequence of situational and transformational thoughts. *Transformations* cover such changes as negating, pluralizing, doubling, taking away the number you first thought of—any class of move away from the original situation which renders the post-transformational situation determinate. Skilled reproductive thinking requires the development of *transformation-selecting* habits as well as *transformation-applying* habits. For example, Bartlett's problem (to fill the gap in ' 1, , 17 ') requires the selection of a transformation and then its application.

Like overt behaviour chains, chains of transformational responses may be functionally equivalent (compare ' double, then add 2 ' with ' add 1, then double '), and so the conditions for the development of transformational habit-family hierarchies are met. As the number of transformational chains established in a given area increases, the repertoire will begin to display increasingly complex characteristics, with overlapping and cross-cutting transformational habit-families. When all possible chains have been established, the structure of the repertoire may resemble that of a mathematical group (meeting the requirements of closure, associativity, identity and inverse). A repertoire with these characteristics, which appear to be necessary for formal operational thinking, is termed a *transitive-group habit-family hierarchy*. By now, however, this elastic concept is considerably overstretched, for the repertoire contains many habit-families (n^2 when there are n possible post-transformational situations), and, since all member transformations are equally accessible, they do not form a hierarchy !

This bald and incomplete summary does scant justice to the care with which each step in the argument is elaborated, and the wealth of reference included by way of substantiation. Ideas from information theorists and computer simulationists, and from Pavlovians of all generations, are skilfully woven into the fabric of the whole. A lucid critique of Piaget's views on conservation, reversibility and equilibrium is particularly welcome

Three chapters on motivational aspects restate Berlyne's belief in conceptual conflict as a powerful stimulus to intellectual activity Thanks largely to Russian researches in educational psychology, a convincing case can be presented, and some of its implications for teaching are indicated. Their supporters (and opponents) will find here an excellent rationale for the use of discovery methods.

William P. Brown.

Davis, F. B. (1965). *Educational Measurements and their Interpretation.* California : Wadsworth, pp. 422, 60s.

It might be thought unfair to judge this book by standards which we in this country hold to be valid. After all these standards arose from the work of such giants as Spearman, Thomson and Burt, and were consolidated by people standing on their shoulders such that someone like Vernon sees further than they did. Such standards of judgment are obviously too severe, and might also be considered irrelevant. The book was, after all, written with the American market in mind: a market having a very big standard deviation Are there relevant American standards for judging it · standards to which we would be willing to subscribe ?

In my opinion there are. They derive from the Chicago School; from the work of Tyler, Bloom, Furst and others. Furst* has stated them most succinctly. They are that work in the field of educational measurement should help people to:

(*a*) Acquire an understanding of the several basic problems involved in developing any evaluation technique.

(*b*) Further their understanding of principles of test construction.

(*c*) Become familiar with some of the best references in the field.

(*d*) Broaden their understanding of the purposes which can be served by tests and other techniques of appraisal.

On all except (*b*) this book falls down and even on (*b*) it does not do the job well. The author set out to provide an introduction to educational measurement for practicing teachers and he should be complimented for seeing a need. For American teachers who come at the subject from a different background of experience, he may have achieved a partial success. For English teachers, the book cannot be recommended.

P. H. TAYLOR.

* FURST, E. J. (1958). *Constructing Evaluation Instruments.* Longmans.

EYSENCK, H. J. (Ed.) (1964). *Experiments in Motivation.* Oxford: Pergamon Press, pp. viii+424, £5.

Most readers of this *Journal* are likely to be familiar with Professor Eysenck's considerable, if controversial, contributions to the study of personality. In this book, however, he " has turned aside from work in this field to study the mysterious but obviously important effects of motivation or drive on behaviour." It is not a book in the ordinary sense, but rather a collection of twenty-four papers, all of them accounts of researches carried out by Eysenck and his associates at the Maudsley Hospital. These are divided into two parts, nineteen papers reporting and discussing experiments with human subjects, the remaining five experiments with animal subjects. Both parts have brief introductions, which discuss the purposes of the later experiments and the theoretical contexts from which they arise. Finally, there is a long bibliography of some twenty-nine pages and a subject index.

Inevitably, this volume must mainly be of interest to readers with advanced and specialised psychological knowledge. No attempt is made by its editor to provide any general conclusions or summing up, or to hint at possible applications of the results obtained and of the theories used. Although the two introductions are admirably lucid, the accounts which follow of the experiments performed are often highly technical; sometimes, indeed, they are overburdened with detail to the point of tedium and triviality. One cannot help wondering whether some of the papers, or parts of them, appearing in this book, many of which are based on Ph.D. theses, might not more profitably have been published in appropriate specialist journals. All in all, this is not a book for the general reader.

Nevertheless, careful and competent work of some interest and importance to experimental psychology is to be found in this volume. Most of the experiments with human subjects reported in Part I compare the performances on a variety of tasks of " candidates for entrance to the apprentice training school of a large automobile manufacturer " with the performances on the same tasks of " students already attending the school." As admission to the training school is highly competitive, only about 10 per cent. of those applying gaining places, the plausible assumption is made that the first or ' high drive group ' is much more powerfully motivated to perform the tasks successfully than is the second or ' low drive group'. The experiments with animal subjects reported in Part II compare the often different behaviour patterns of rats specially selectively bred for ' high or low emotionality '; the analogy with the work on humans is that the emotional rats may be thought of as having high drive and the non-emotional rats as having low drive. A clear account of the reasons for, and methods of, distinguishing between these two kinds of rats is given in the introduction to Part II by Eysenck and Broadhurst. Continual efforts to integrate the experimental findings reported into the conceptual framework of Hull's behaviour theory are characteristic of both parts of the book, but it is frankly admitted that sometimes this cannot be done and that occasionally the obtained results are more than a little puzzling. In fact, one of this book's charms is its modesty and intellectual honesty. Some, perhaps, will be surprised, if delighted to find Eysenck writing in his foreword : " There is a certain degree of congruence between theories and results, sufficient to encourage one to continue in this type of work, insufficient to make one feel that issues have been clarified even in the most rough and ready manner."

As usual, Pergamon Press has managed to produce a prohibitively expensive book. This we have come to expect. What is even more disturbing is the large number of misprints and other technical deficiences, which are unhappily littered throughout this volume. It would, however, be a pity if justified indignation about both price and shocking technical standards were to prevent readers from recommending this work's purchase by their departmental or college library.

G. W. PILKINGTON.

HORROCKS, J. H. (1965). *Assessment of Behavior.* Columbus, Ohio : Charles E. Merrill Books, Inc., pp. xv+736, 68s.

A sound though somewhat unexciting book ! It contains all the student needs to know about selected intelligence, aptitude, achievement and personality tests, with useful introductory chapters and sections on background topics such as mental measurement, reliability, the nature of intelligence and theories of personality. Dr. Horrocks specially stresses the need for students to understand what is involved in the Assessment of Behavior, as well as possessing information about measuring instruments. To the reviewer, however, the book is not particularly helpful in this respect ; the theoretical discussions show nothing like the insight, for example, of Cronbach's ' Essentials,' and are sometimes rather muddled. What, for instance, is one to make of the following statement about approaches to personality ?

Elementarianism stems in part from the traditional British position that human nature is susceptible to analysis on the basis of three relatively separate components : cognitive, conative and emotional. Individuals are seen as differing in personalities in proportion to the qualitative and quantitative degrees to which these three components are manifested in their behavior.

Again, many theoretical issues which are of vital importance in test interpretation are neglected or inadequately explained, e.g., construct validity, variation of I.Q. standard deviations, practice and coaching effects, age changes in the Wechsler, why differential aptitude or factor tests don't differentiate, response sets in attitude questionnaires, and so on. No ! The strength of the book lies in the detailed information it gives about the materials and administration of particular tests, how they were standardised and what is known of their reliability and validity. The historical surveys of early work on the development of different areas of testing are good also, and the author is scrupulously fair in setting out the views of critics as well as supporters (e.g., of projective techniques).

There are several similar books already published, and the reviewer would prefer to use Cronbach, Anastasi or Freeman for course work with British students (partly because they are less verbose). As Horricks deals in detail with fewer tests, it would be best to indicate here the major inclusions and exclusions.

Ch. 1–4. Concepts of measurement in psychology ; distributions and elementary statistical concepts, but not techniques ; reliability, validity ; classification of different types of tests for different purposes.

Ch. 5. Nature of intelligence ; elementary factorial principles and findings.
Ch. 6. The Binet scales, in great detail.
Ch. 7. Performance tests.
Ch. 8. The Wechsler scales.
Ch. 9. Group tests (almost wholly American).
Ch. 10. ' Culture-free ' tests, including Raven.

Ch. 11–12. Factor classifications of French, Guilford, Ahmavaara. Special aptitude tests and their construction. Various mechanical and clerical tests; Bennett's and some other batteries of differential tests.

Ch. 13–14. Physical-motor development; Gesell, P. Cattell, Minnesota, Griffiths; also reading and other readiness tests. Seashore and Wing music tests.

Ch. 15. Achievement tests and their purposes in education. Some discussion of essay *versus* objective examinations and their construction; but omits descriptions of particular tests or batteries.

Ch. 16. Personality, definitions and theories (too briefy to be helpful).

Ch. 17. Inventories: Edwards PPS, MMPI, California Test of Personality, Guilford's and IPAT scales.

Ch. 18. Personality ratings (a thorough discussion).

Ch. 19. Projective techniques: Rorschach, TAT and its derivatives, Rosenzweig PF and others, but omitting drawing and doll-play.

Ch. 20. Interests, attitudes, social behaviour; Strong and Kuder; Thurstone attitude scales, Allport-Vernon-Lindzey; Sociometry; but no discussion of objective or observational techniques, nor of the interview.

There are a few misspellings of authors: the following caught the reviewer's eye: D. G. Paterson (pp. 169, 188); W. McDougall (pp. 492, 520, 727); J. McK. Cattell (p. 601). P. E. VERNON.

JOHNSON, R. C., and MEDINNUS, G. (1965). *Child Psychology: Behavior and Development.* New York: John Wiley, pp. x+556, 53s.

The taxonomic problem with omnibus books on child psychology is still unsolved. With the increasing emphasis on trying to explain how and why children change, the division into various aspects of development that succeeded normative, age-stage descriptions is now being superseded by using the main formative influences as chapter headings. The emphasis, as in Johnson and Medinnus' book, being mainly on socialisation and personality, there is less fragmentation of the child, but an inevitable tendency either to talk about ' children ' as a group, or to oscillate between stages of development.

Child Psychology has five sections. The first section, a detailed account of methods and techniques used in child psychology, would be admirable if less undiscriminating, e.g., a dated quotation from Stanley Hall occupies two pages, reliability and validity get two paragraphs. This is a recurring fault. The rest of the book aims to show how people are similar and different as a result of two main forces—' biological-genetic,' mainly producing similarities, and social, mainly producing differences. In the second section, on areas seen as largely responsible for the common quality of human beings like heredity, maturation and learning, the emphasis, especially and surprisingly in the chapter on intelligence, is deliberately on biological-hereditary factors. Old friends such as the Hopi Indians and Albert B. make their predictable appearance accompanied by many new examples which are given in considerably greater detail than the theoretical explanations they are supposed to illuminate: this makes the book interesting but difficult or misleading to a beginner. The biggest weakness of the whole book, however, is that these chapters contain nearly everything that the authors have to say about cognitive development, which is consequently quite inadequately covered. For example, concept formation occupies a page, and the few cursory references to Piaget deal more with his early work than with his stages of intellectual behaviour.

The next two sections, dealing respectively with the family and other societal influences, are the best part of the book. The chapters on the influence of the family and peer relationships are especially good: wherever possible the best available evidence is quoted, and when speculation appears this is clearly stated.

There follows a standard chapter on intelligence and personality tests, a disappointingly abstract chapter on behaviour problems, and a good chapter on personality development that recapitulates much previous material together with a summary of research findings on some major aspects of personality.

This book is best suited to a psychology student who wants to add a thorough understanding of the societal influences on children's personality to his knowledge. It cannot, however, be recommended as a sole text on child psychology to prospective teachers.

RALPH D. DUTCH.

KAHN, J. H. (1965). *Human Growth, The Development of Personality*. Oxford: Pergamon Press, pp. xvi+220, 25s.

In this book of just over 200 pages, the author attempts to survey the whole field of human growth from birth through childhood and adolescence to middle and old age. He deals with human needs, definitions of personality and developmental stages in the early part of the book; giving a brief but interesting survey of these aspects of the subject. This is followed by a series of chapters entitled 'The Suckling,' 'Toilet Training' and 'Infantile Sexuality' which are psycho-analytic in terminology and in explanation of the forces affecting child development; as are the chapters on 'Rejection' and 'Acceptance.' The chapters on education are more general in approach, and many problems of importance at the present time are high-lighted in the few sentences devoted to them. The physical changes and the emotional ones associated with them are considered in the section on adolescence, adult life, middle and old age. Brief mention is also made of the more typical mental illnesses, within the framework of the physical changes that are taking place at each stage. Throughout the book the problems considered are discussed within the context of their social setting.

As claimed in the foreword, this book should be of value to students and to general reader. I would not regard it as in any sense a textbook, but merely as background reading to be supplemented in at least some of the areas. Any author who attempts within a single book of this size to cover so extensive a field must of necessity sacrifice depth in any given aspect. Lectures to social workers, probation officers and nurses form the basis of this book. In this may lie one of its deficiences. What may be suitable for oral presentation may not lend itself to written presentation without a great deal of amendment. Drastic cutting in certain parts, and re-organisation of others, would have improved the book; while an index and a glossary would have been useful additions.

MARGARET M. CLARK.

MATHER, D. R., FRANCE, N., and SARE, G. T. (1965). *The C.S.E.—A Handbook for Moderators*. London: Collins, pp. 188, 15s.

The introduction of the new C.S.E. examination, in some areas in 1965, in all areas by 1966, is bound to lead to increased activity among the educational publishers. No doubt many of the resulting text books will be of the kind that most educationists will deplore: cram books designed to outwit the examiner rather than to educate the pupil. All the more reason for schools to opt for 'Mode 3' examining—internal examination externally moderated—rather than 'Mode 1'—external examination. The book under review is not one of those whose publication should be deprecated: quite the reverse. One of the major problems associated with the launching of a new examination, and particularly of one expressly designed to be teacher-controlled, is the education and training of teachers in up-to-date techniques of examining, and the associated skills such as item-analysis and elementary statistical method. The Department of Education and Science, through its Curriculum Study Group (now merged into the Schools Council), was keenly aware of this from the first moment of the birth of C.S.E., and it has sponsored the publication of a series of Bulletins (five have been issued up to now) which have been planned to form the basis for such education and training. These are excellent publications, and it is hoped that they will find a wide circulation in our colleges of education as well as among serving teachers: the relative paucity of sound courses on the philosophy and techniques of examining in initial training is deplorable. These Bulletins, however, give only a brief and rather sketchy introduction to the field, and it is essential that they should be buttressed by fuller and more extensive treatments Existing texts are nearly all American, and these have obvious limitations for the British teacher, with their excessive emphasis on multiple-choice examinations and their neglect of the essay-type answer (with all its advantages and with its many problems) This book is a welcome addition to the small number of British texts in this field

The book is in two parts. The first part gives a history of the C.S.E. and its development from the Beloe Report. One might criticize some of the detail included here—such as extensive quotation from speeches and reports—but the story of the successful emergence of the Curriculum Study Group through a fog of initial mistrust and misunderstanding, and its final incorporation into the Schools Council, is a fascinating one, and one which all teachers could study with profit. The G.C.E. comes in for some hard criticism, mainly in order to emphasise the dangers which the C S.E must try to avoid.

The second part of the book is that which justifies the book's sub-title, " A Handbook for Moderators " It covers questions of reliability and validity, of item-analysis, of grading and the fixing of pass-marks, and it tackles the question of the moderator's role and the techniques by which he can pursue his aims and purposes. This section builds on the Examinations Bulletins already mentioned, and adopts the ' short-cut ' methods of Bulletin 5—those admirable techniques devised by Mr. Gilbert Peaker by which *range* is used instead of standard deviation in calculating tolerances and significances. The only criticisms one has of this part of the book are minor—the lack of emphasis, for example, of the ' goal-oriented ' view of examinations, rather than the ' syllabus-sampling ' approach ; the recommendation of 100 candidates for test try-out, when 150+ would be better, and the suggestion (p. 81) that a rectangular distribution can be obtained by " spreading evenly the difficulty of the questions " The last point is a constantly recurring one in text-books, in spite of the evidence provided on both sides of the Atlantic that the choice of questions as near to 50 per cent. difficulty as possible offers the best chance of flattening the distribution curve. But these are minor points in a book which can be recommended for all teachers concerned with C.S.E.—and for all education lecturers (and subject lecturers) in colleges of education. STEPHEN WISEMAN.

WILKINSON, A., with contributions by DAVIES, A., and ATKINSON, D. (1965). *Spoken English*. Birmingham : Educational Review, Occasional Publication No. 2, pp. 130, no price.

This book is urgently needed. Its aim is to put Spoken English on the map in education. Its authors are a team working on listening comprehension at Birmingham University ; they have a thorough knowledge of what goes on in schools and are concerned to make information about recent work on Spoken English readily available to teachers and trainers of teachers. It is the sense that there is matter here to be *acted* upon, and that action needs to be based on objective work—in what is a myth-prone area—which gives the book its force The authors say of *The Linguistic Sciences and Language Teaching*, by Halliday, McIntosh and Strevens (1964), " if only sufficient *teachers* of English can be persuaded beyond the formidable title, it could itself reshape language attitudes and language teaching in British education " This might well be said of this book, which has no such formidable title and whose demands on its readers are much more permissive

It seems, at first sight, to consist of an odd assortment of items. There are eight articles, more or less loosely connected by the fact that they all deal with aspects of Spoken English. The first part of the book examines the paradoxical situation in which Spoken English is badly neglected in schools while it is a central area of research in linguistics, sociology and the psychology of language. The second part is concerned with tests of the Spoken Language ; it uses the concepts of linguistics to disperse the preconceived ideas that govern most of our attitudes to speech, and to provide an analytical framework for testing it, and for the criteria for these tests. It is, perhaps, the last item above all—" a dummy test of listening comprehension, for demonstration purposes only " which catapults the reader into those new areas which workers in the field of Spoken English are most concerned to bring to his attention—the Spoken Language as it really is. While the actual matter of the book is easy and forthright reading, its foundations are solid ; it is fully documented with research references throughout, and these and the appendices provide a bibliography that would cover serious study in any of the areas touched on in the book.

Its disconnectedness is, perhaps, that of good talk. The authors offer their observations " as an early contribution to a *debate* on Spoken English " and the book itself has some of the strengths and weaknesses of a well-informed discussion. There are some gaps and occasional misleadings. Why for instance, in a book of this calibre, is not the question explored of what is being tested besides the candidate ? In the earlier part of the book, Dr. Wilkinson is at pains to show that " speech and personality are one " and that " certain patterns of social behaviour are more favourable to oracy than others." It this is so, then personality and social background are also being tested and these cannot be included in "systematic language teaching " in the classroom. This, of course, is the argument behind the objections to fostering tests of Spoken English as part of our national examinations system There are arguments on both sides and since they are hotly debated by teachers, it is a pity they are not explicitly considered here.

Again, Alan Davies makes specific reference to a scheme for systematic language teaching in another book. He omits to mention that this scheme is presented in the context of teaching a foreign language. Children know their mother tongue before they come to school so inevitably there must be different aims and procedures. To suggest, by implication, direct application to teaching the mother tongue is misleading.

But those matters spring from practical needs. Oral examinations are with us, and the need for more systematic teaching of the mother tongue is clear—though tests and systems of themselves will not improve things, and, in fact, the tenour of the whole book is to provide *theoretical* backing for necessary changes in attitude and practice. As the authors say, " we are really at the beginnings of serious study in many cases." This book is a very good beginning. N. C. MARTIN.

PUBLICATIONS RECEIVED

The mention of a book in this list neither implies nor precludes a later review.

Joan Wynn Reeves (1965). *Thinking About Thinking.* London: Secker and Warburg, pp. 333, 55s.

Matti Koskenniemi (1965). *The Development of Young Elementary School Teachers.* Helsinki: The Finnish Academy of Science and Letters, pp. 635, no price.

Demos and Grant (1965). *Vocational Guidance Reading.* Illinois: Charles C. Thomas, pp. $15.50.

Cyril Burt (1965). *Studies in Psychology.* London: University of London Press, Ltd., pp. 283, 42.

M. D. Vernon (1965). *The Psychology of Perception.* London: University of London Press, Ltd., pp. 235, 25s.

J. C. Eccles, L. Kruger, H. Hyden, J. McGaugh, A. M. Uttley, H. von Foerster, D. P. Kimble (1965). *The Anatomy of Memory, Vol I.* U.S.A.: Science and Behavior Books, Inc., pp. xi+451, no price.

Boletim Bibliografico (1965). Centro de Investigacao Pedagogica Fundacao Calouste Gulbenkian, pp. 289, no price.

Robert F. Mager (1965). *Preparing Instructional Objectives.* U.S.A.: Fearon Publishers, pp. x+62, no price.

Esi Quarterly Report (1965). U.S.A.: Educational Services Incorporated, Massachusetts, pp. 267, no price.

Udai Narain Pareek (1965). *Developmental Patterns in Reactions to Frustration.* London: Asia Publishing House, pp. vii+182, 25s.

Warren Lamb (1965). *Posture and Gesture: An Introduction to the Study of Physical Behaviour.* London: Gerald Duckworth and Co., pp. 189, 25s.

W. A. L. Blyth (1965). *English Primary Education,* Vol. I. London: Routledge and Kegan Paul, pp. xv+214, 25s.

W. A. L. Blyth (1965). *English Primary Education,* Vol. II. London: Routledge and Kegan Paul, pp. xi+156, 25s.

F. H. George (1965). *Basic Principles of Digital Computing.* Bristol: Teaching Programmes, Ltd., pp. 43, 7s. 6d.

Henry W. Maier (1965). *Three Theories of Child Development.* New York: Harper and Row, pp. xvi+314, 30s.

Oscar Krisen Buros (1965). *The Sixth Mental Measurements Yearbook.* New Jersey: The Gryphon Press, pp. xxxvi+1200+B526+index, $32.50.

R. V. Sampson (1965). *Equality and Power.* London: Heinemann, pp. 247, 35s.

Margaret Mead (1965). *Anthropologists and What They Do.* New York: Franklin Watts, Inc., pp. 209, $3.95.

Mary D. Sheridan (1965). *The Handicapped Child and His Home.* London: National Children's Home, pp. 63, 7s. 6d.

Paul Halmos (1965). *The Faith of the Councillors.* London: Constable and Co., pp. 220, 30s.

Norman L. Munn (1965). *The Evolution and Growth of Human Behavior.* London: George Harrap, pp. xi+594, 48s.

Gardner Lindzey and Calvin S. Hall (1965). *Theories of Personality: Primary Sources and Research.* New York: John Wiley, pp. xii+543, 53s.

Joseph Majault (1965). *Teacher Training.* Strasbourg: Council for Cultural Co-operation of the Council of Europe, pp. 211, 8s. 6d.

E. Graham Howe (1965). *Cure or Heal.* London: George Allen and Unwin, pp. 236, 28s.

J. Zubin, L. D. Eron and F. Schumer (1965). *An Experimental Approach to Projective Techniques.* New York: John Wiley, pp. xix+645, 105s.

B. Grant, G. D. Demos and W. Edwards (1965). *Guidance for Youth.* U.S.A: George C. Thomas, pp. xiii+113, $4.50.

M. A. Wallach and N. Kogan (1965). *Modes of Thinking in Young Children.* New York : Holt, Rinehart and Winston, Inc., pp. viii+357, $8.

A. E. Goss and C. F. Nodine (1965). *Paired-Associates Learning—The Role of Meaningfulness, Similarity and Familiarization.* New York : Academic Press, pp. ix+358, 108s.

M. Craft (1965). *Ten Studies into Psychopathic Personality.* Bristol : John Wright, pp. 136, 25s.

Walter B. Barbe (1965). *Teaching Reading* : *Selected Materials.* New York : Oxford University Press, pp. xiii+444, 32s.

Don C. Dinkmeyer (1965). *Child Development—The Emerging Self.* New Jersey : Prentice-Hall, pp. xi+434, no price.

Wayne, Dennis (1965). *The Hopi Child.* New York : John Wiley, pp. vii+200, $1.45.

Carl I. Hovland, Arthur A. Lumsdaine and Fred D. Sheffield (1965). *Experiments on Mass Communication.* New York : John Wiley, pp. x+345, 15s.

J. F. Eggleston (1965). *A Critical Review of Assessment Procedures in Secondary School Science.* School of Education University of Leicester, pp. 55, 5s.

Hans W. Bentz (1961). *Sigmund W. Freud in Ubersetzungen.* Frankfurt : Hans W. Bentz, pp. 60, no price.

(1965). *Paedagogica Europaea* : The European Yearbook of Educational Research. Amsterdam and Brussels : Agon Elsevier ; Braunschweig : Georg Westermann Verlag, pp. xii+284, 75s.

D. W. Winnicott (1965). *The Maturational Progresses and the Facilitating Environment.* London : Hogarth Press, pp. 295, 42s.

Martin Buber (1965). *The Knowledge of Man.* George Allen and Unwin, Ltd., pp. 186, 28s. net.

William Kessen (1965). *The Child.* New York : John Wiley, pp. xi+301, 38s. cloth, 23s. paper.

(1965). *The Attainments of Scottish Ten-year-olds in English and Arithmetic.* London : University of London Press, Ltd., 94, 5s.

Robert L. Ebel (1965). *Measuring Educational Achievement.* New Jersey : Prentice-Hall, Inc., pp. xii+481, 64s.

George A. Talland (1965). *Deranged Memory.* New York : Academic Press, pp. xi+356, 76s.

Richard R. Skemp (1965). *Understanding Mathematics, Book* 2. London : University of London Press, Ltd., pp. 358, no price.

C. Williams (1965). *Graduation and Employment.* Queensland : University of Queensland Press, pp. 40, 9s.

R. W. J. Keeble (1965). *A Life Full of Meaning.* Oxford : Pergamon Press, pp. xi+226, 21s.

Vivian Gourevitch (1965). *Statistical Methods* : *A Problem-solving Approach.* Boston : Allynand Bacon, Inc., pp. x+310, $7.75.

L. P. Lipsitt, Charles C. Spiker (1965). *Advances in Child Development and Behavior.* New York : Academic Press, pp. x+269, $8.50.

E. H. Schein (1965). *Personal and Organizational Change through Group Methods.* New York : John Wiley, pp. x+375, 62s.

S. H. Irvine (1965). *Selection for Secondary Education in Southern Rhodesia.* Salisbury : University College of Rhodesia and Nyasaland, pp. vii+35, 7s. 6d.

Editorial Committee and Contributing Librarians (1965). *Mental Health Book Review Index.* New York : Council on Research in Bibliography, Inc., pp. 78, no price.

Robert S. Woodworth (1965). *Contemporary Schools of Psychology.* London : Methuen and Co., pp. vii+455, 45s.

G. A Ferguson (1965). *Nonparametric Trend Analysis.* Montreal : McGill University Press, pp. v+61, $1.95.

EDUCATION
TODAY

A paperback series which covers the full range of educational topics of current interest and development. The series includes practical handbooks for classroom use as well as theoretical discussion of controversial subjects, at all levels from the nursery school to further education.

INVESTMENT IN CHILDREN

Editor: M. L. Kellmer Pringle

This symposium on positive child care and constructive education develops ideas presented at the conference organised by the National Bureau for Co-operation in Child Care on the subject: Prevention—What and How?

Just published. 12s. 6d. net.

STUDENT GUIDANCE

F. Claude Palmer

'A stock of valuable and thoroughly sound ideas about . . . a new and important job in the educational scene.' TIMES EDUCATIONAL SUPPLEMENT. *10s. net.*

PROGRAMMED LEARNING IN THE SCHOOLS

J. Leedham and D. Unwin

'A balanced survey of what has been achieved, and what may develop . . . by two people of experience in the field. A broad sweep of background information, discriminating advice and reasoned judgment which can be unreservedly applauded.'
TIMES EDUCATIONAL SUPPLEMENT. *10s. 6d. net.*

SOCIAL LEARNING AND ITS MEASUREMENT

M. L. Kellmer Pringle

The practical implications of social learning have been, and still are being, neglected. This book both reviews the development and role of social learning, and reports the main findings of a pilot study conducted by the author.

Ready April. Probably 12s. 6d. net.

1724

Longmans

THE BRITISH JOURNAL OF MEDICAL PSYCHOLOGY

Edited by Dr. THOMAS FREEMAN

Volume 39, *Part* 1. CONTENTS *March,* 1966

Eugene Wolf. Learning Theory and Psycho-Analysis.
I. M. Marks and M. G. Gelder. Common Ground Between Behaviour Therapy and Psychodynamic Methods.
John E. Gedo. The Psychotherapy of Developmental Arrest.
Adah Maurer. Maturation of Concepts of Death.
Pinches Noy and Shlomoh Wollstein. Clinical Observations on the Psychogenesis of Impotence.
Alec J. Coppen. The Marke-Nyman Temperament Scale : An English Translation.
Harvey R. Greenberg. Notes on the Parental Exclusion Phenomenon in Twins.
Richard M. Suinn. Identification and Authoritarianism.
Critical Notice :
Reviews :

Subscription £4 10s. per volume of 4 parts. Single parts 30s. plus postage

Orders may be placed with any bookseller, or direct to :

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200, Euston Road, London, N.W.1

THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY ARTHUR SUMMERFIELD

CONTENTS OF VOLUME 56, PART 4, 1965

Neisser, Ulric and Beller, Henry K. Searching through word lists.
Neisser, Ulric and Stoper, Arnold Redirecting the search process.
Treisman, Anne M. The effects of redundancy and familiarity on translating and repeating back a foreign and a native language.
Furth, Hans G. and Youniss, James. The influence of language and experience on discovery and use of logical symbols.
Matthew, W. A. The relation between association norms and word frequency.
Brown, W. P. Emotional indicators in word association.
Murdock, Bennet B , Jr. Effects of a subsidiary task on short-term memory.
Pickford, R. W. Two artists with protan colour vision defects.
Corteen, R. S and Blackman, A. R. Skin conductance change and sensory discrimination.
McFarland, Joseph H. Sequential part presentation . a method studying visual form perception.
Warrington, Elizabeth K. The effect of stimulus configuration on the incidence of the completion phenomenon.
Hermelin, B. and O'Connor, N. Visual imperception in psychotic children.
Schwartz, Stewart and Spitzer, Gilbert. Preferred handedness as a factor in the ability to form a set based on imaginal situations.
Murch, Gerald M. A simple laboratory demonstration of subception
Orme, J. E. Ability and season of birth.
Arthur, A. Z. 'Cross-over' effect and subjective intensity.
Audley, R. J. and Wallis, C P. 'Cross-over' effect ; a reply to Arthur.
McPherson, Frank M Comment on Eysenck's account of some Aberdeen studies of introversion-extraversion and eyeblink conditioning.
Publications Reviewed
Other Publications Received

Subscriptions 90s. net per volume of 4 parts
Orders may be placed with a bookseller or sent direct to the publishers

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London, N.W.1

OCCUPATIONAL PSYCHOLOGY

Editor : ALEC RODGER

JULY 1965 VOLUME 39, No. 3

CONTENTS

Intervening Responses to Changes in Supervisor Job Designs.
By LOUIS E. DAVIS and ERNST S. VALFER

"... All Except ..." Definition by Exclusion, Relative Class Complexity and S—R Compatibility in Dichotomous Classification Tasks.
By R. D. SHEPHERD

A Leap Forward in the Social Sciences ?
By GEORGE WESTBY

The Assessment of Undergraduates by their Teachers.
By D. J. WARREN PIPER

Vocational Intentions and Decisions of Aberdeen Arts Graduates.
By J. D. NISBET and WILMA GRANT

Book Reviews

Annual Subscription 40 *shillings, overseas* 50 *shillings*

National Institute of Industrial Psychology

14, Welbeck Street, London, W.1.

AUSTRALIAN JOURNAL OF PSYCHOLOGY

Editor : I. K. WATERHOUSE, University of Melbourne

Vol. 17 No. 3 *December*, 1965

KNOPFELMACHER, F. Types of Behaviour Theory.

FEATHER, N. T. Reactions to Communications under Conditions of Source Responsibility and Source Coercion.

GREGSON, R. A. M. Representation of Taste Mixture Cross-modal Matching in a Minkowski r – metric.

FUREDY, J. J. Reinforcement through UCS Offset in Classical Aversive Conditioning.

NAYLOR, G. F. K. Subjective Judgements of the Vertical: and Estimate of their Nature and Consistency.

POLLARD, J. S., LYSONS, A. M. and HUGHES, R. N. Size of Closed Field Test and Performance of Rats.

MCLEOD, J A Comparison of WISC Sub-Test Scores of Pre-Adolescent Successful and Unsuccessful Readers.

KRAUS, J. Cattell Anxiety Scales Scores and WAIS Attainment in Three Groups of Psychiatric Patients.

Book Reviews

The Journal is published three times a year. Subscription rates 36s. (sterling), and $5·50 in U.S.A. Members of B.P.S. and A.P.A. should enquire about special subscription rates from the Secretary of their Society. Non-members should send enquiries to Melbourne University Press, Parkville, N.2., Victoria, Australia, or to its agents.

in U.K.
Cambridge University Press,
200 Euston Road,
LONDON, N.W.1

in U.S.A.
Cambridge University Press,
32 East 57th Street,
NEW YORK. 22, N.Y.

Enquiries from the Orient should be sent directly to Melbourne University Press.

Methuen's Manuals of Modern Psychology

THE PSYCHOLOGY OF CHILDHOOD AND ADOLESCENCE

C. I. Sandström

An introductory book for students on developmental psychology. It describes the psychology of childhood from the pre-natal stages up to adolescence and beyond. C. A. Mace contributes a special introduction.

With 32 plates and 55 line drawings, **36s.**

IMPRINTING AND EARLY LEARNING

W. Sluckin

The importance of imprinting observed in young animals is widely recognized to-day by students of behaviour. In this book the author traces the history of the interest in this subject and describes the work done so far.

With 9 diagrams, **25s.**

METHUEN *& Co Ltd · 11 New Fetter Lane · London EC4*

ELEMENTARY STATISTICAL METHODS IN PSYCHOLOGY AND EDUCATION

Paul Blommers and E. F. Lindquist

Statistical methods are the techniques used to interpret collections of quantitative or numerical data. The examples used in this book have been drawn mainly from the fields of psychology and education. The main emphasis is on developing a knowledge of the appropriate technique to select for a given purpose and on a critical interpretation of results. This book will enable the student to develop a highly critical attitude towards over-generalisation and stereotyped evaluation. *Publishing March* 45*s. net*

Study Manual *Publishing March* 10*s.* 6*d. net*

Order from your local bookseller.

UNIVERSITY OF LONDON PRESS LTD

St. Paul's House
Warwick Lane
London EC4

The British Journal of

Social and Clinical Psychology

Edited by MICHAEL ARGYLE and R. R. HETHERINGTON

Contents of Vol. 5, Part 1, February, 1966

JAHODA, G. Impressions of Nationalities—an Alternative to the 'Stereotype' Approach.

MORRISON, A. and McINTYRE, D. The Attitudes of Students towards International Affairs.

MURDOCH, P. H. J. Birth Order and Age at Marriage.

RIM, Y. Machiavellianism and Decisions Involving Risk.

WILDE, G. J. S. The Significance of Decided and Query Answers to Lie-Items in a Personality Questionnaire.

BARBER, T. X. The Effects of 'Hypnosis' and Motivational Suggestions on Strength and Endurance: A Critical Review of Research Studies.

FRANSELLA, FAY and ADAMS, B. An Illustration of the Use of Repertory Grid Technique in a Clinical Setting.

CLARIDGE, G. S., WAWMAN, R. J. and DAVIES, M. H. Sedation Threshold, Spiral After-effect and Overinclusion.

BOOK REVIEWS

The subscription price per volume of four parts is 90s. net. Orders may be sent to your bookseller or direct to the publishers

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200, Euston Road, London, N.W.1

THE
BIRMINGHAM
PRINTERS LTD
PRINTERS & BOOKBINDERS

High Class
Technical
Publications

42-44 HILL STREET
BIRMINGHAM 5
TELEPHONE: MIDLAND 0196-7

VOLUME XXXVI JUNE 1966 PART 2

THE SOCIAL NEEDS AND SATISFACTIONS OF SOME YOUNG PEOPLE

PART II—AT SCHOOL

BY F. MUSGROVE
(University of Bradford)

V.—SCHOOL.

Whereas home, youth clubs and work appeared substantially to meet the needs of young people in a Northern conurbation, their schools did not. By far the highest proportion of negative statements were made on a sentence-completion schedule about school, particularly about the grammar school. 19·5 per cent. of secondary modern school pupils' statements were negative; 32·8 per cent. of grammar school pupils' statements were negative. (Pupils below the sixth form made 28·0 per cent. negative statements, sixth formers 37·3 per cent.). By contrast, 9·7 per cent. of the statements about home were negative, 5·9 per cent. of statements about the club, and 7·3 per cent. of statements about work.

In summary, the demands of modern school pupils were: Expressive (E) 57·1 per cent., Instrumental (I) 42·9 per cent. Their satisfactions were: Expressive, positive (E+): 32·4 per cent., Expressive, negative (E –): 19·5 per cent., and Instrumental (I): 48·1 per cent. The demands of grammar school pupils were: E 45·6 per cent., I 54·4 per cent. Their satisfactions were: E+ 23·6 per cent., E –32·8 per cent., and I 43·6 per cent.

School is seen in an ' instrumental ' light to a far greater extent than home or club. Grammar school pupils want it to be far more instrumental than modern school pupils: 54·4 per cent. of their demand is for instrumentality (chiefly in the sense of intellectual training, preparing for examinations, and helping you to get on in life) compared with 42·9 per cent. of the modern school pupils' demand (CR 3·1, $P < 0{\cdot}01$). But whereas grammar school pupils mention instrumental satisfactions less frequently than they mention instrumental expectations (43·6 per cent. compared with 54·4 per cent.), modern school pupils mention them more often (48·1 per cent. compared with 42·9 per cent.). Grammar school pupils appear to get less instrumentality than they desire, modern school pupils appear to get a surfeit.

Intellectual training and preparation for a career overshadow all other instrumental demands and satisfactions. It is not only the grammar school pupils who want to learn. Thus, a 15-year-old modern school girl, daughter of a warehouseman, says: "At school you always have plenty of chance to have things explained which you don't understand." More commonly, they make general statements to the effect that at school you have plenty of chance to learn and ' get a good education.' The opportunity to develop physical skills is mentioned much less frequently by both grammar school and modern school pupils.

TABLE 5

STATEMENTS ABOUT INTELLECTUAL AND PHYSICAL DEVELOPMENT AT SCHOOL.

	Demand	Satisfaction
All boys (N : 336)		
Intellectual development	23·8%	30·9%
Physical development	4·2	4·0
All girls (N : 264)		
Intellectual development	22·7	21·9
Physical development	0·4	1·5

It is in its expressive rather than its instrumental role that school falls far short of demand. The major expressive demand is for freedom and self-direction : freedom to put your own point of view, to be treated as an individual. This is also a major demand of club and home ; but these institutions appear to come nearer than the school to satisfying the demand.

TABLE 6

STATEMENTS ABOUT FREEDOM AND SELF-DIRECTION
(Boys and Girls)

	Demand	Satisfaction	d
School :			
Boys (N : 336)	22·0%	3·6%	—18·4%
Girls (N : 264)	22·3	8·3	—14·0
Club :			
Boys (N : 405)	13·5	10·4	—3·1
Girls (N : 345)	13·3	11·0	—2·3
Home :			
Boys (N : 405)	22·5	13·3	—·92
Girls (N : 345)	23·9	11·3	—12·6

Both modern school and grammar school pupils express their appreciation of school as a social institution. " At school you always feel needed and part of the school," says a 14-year-old grammar school boy. "At school you always have a chance to talk to the staff relatively freely," says a 16-year-old grammar school girl. A 15-year-old modern school girl says that "Above all, school helps you to feel adult " ; the 14-year-old son of a truck driver says that "At school you always feel that you are some one, even if you aren't." Another modern school boy expresses a similar sentiment : "At school," he says, " you always feel that you are somebody." "At school," says a 15-year-old modern school girl, " you always feel that the teachers are interested in you."

Like grammar school pupils, modern school pupils also make negative statements, chiefly about restrictions and personal humiliations. "At school," says the 15-year-old daughter of a boiler-man, " you always feel that you cannot do what you want and that teachers are against you." (At home, the same girl finds the freedom she wants : " You always have plenty of chance to do what you like. You're not told to stop talking, like at school.")

Another 15-year-old girl says that at school " you always feel boxed in, regimented, treated like the second form, and angry at the lack of trust." The 15-year-old daughter of a clerk says of her modern school that " you always feel that you're still a child. They should encourage you to grow up." (She, too, finds freedom at home, where " you always feel loved and highly thought of.") Modern school boys are more likely than grammar school boys to refer to physical humiliations : "At school you always feel that teachers can pull your hair and slap you at the back of the neck."

But grammar school pupils differ significantly from modern school pupils in the frequency of their unfavourable comments about school. It is difficult to explain this simply as a function of higher intelligence and critical powers : grammar school pupils do not exercise their superior intelligence in this way when they make comments on their home, club and work. There is no tendency for working class pupils in grammar schools to make more negative statements than pupils from middle-class homes.

TABLE 7

STATEMENTS OF SATISFACTION AND DISSATISFACTION WITH SCHOOL.
(Boys and Girls)

	' School is'		
	E+	E–	I
Modern School (N : 210)	32·4%	19·5%	48·1%
Grammar School (N : 390)	23·6	32·8	43·6

Chi square 10·18 d.f. 2 P<0·01

There was a tendency for sixth formers to be more negative than pupils below the sixth form, but the difference falls short of significance at the 5 per cent. level. The sixth formers who took part in this inquiry were not borderline academics. The boys had obtained an average of six O-level G.C.E. passes, the girls seven. Some had already distinguished themselves in A-level examinations and were staying on for a third year in the sixth.

TABLE 8

DISSATISFACTIONS OF SIXTH FORMERS.

	E+	E–	I
Sixth Formers (N : 201)	24·4%	37·3%	38·3%
Pupils below thefiSixth (N : 189)	22·8	28·0	49·2

Chi square 5.35 P<0·10.

The great demand is for self-expression and self-direction : " At school you should have plenty of chance to express your views " ; " At school you should always be able to feel that masters are not simply disciplinarians." And one should be recognized and treated as an individual : "At school you should be able to feel that you are a human being and not just a 'B 'pupil " (14-year-old barrister's daughter).

To a large extent, the nine grammar schools contributing to the research population appeared to be failing to meet this demand. "At school you always feel like a little boy " (15-year-old son of a clothing manufacturer) ; At school you always feel disgusted with life " (15-year-old son of a school master) ; "At school you always feel looked down on by masters "(18-year-old son of an advertising agent). The dissatisfaction does not arise from expectations of an easy time which is denied them ; a 17-year-old tailor's son says that at school you should feel expected " to work hard and diligently " ; but he feels " neglected, sick, tired, bored, disgusted and generally miserable."

The 18-year-old son of a branch manager of an insurance company " always feels downgraded " at school ; the 14-year-old son of a manufacturer's agent " always feels a prisoner." The 15-year-old son of a pharmacist finds that at school " you always have a chance to be just one of a conformist rabble : at school you always feel unimportant and hounded." (At home, by contrast, he says he has " plenty of chance to be an individual.") The 15-year-old son of a furniture manufacturer says that " at school you always feel like a chicken which is being stuffed." (He also finds happier circumstances at home, where " you always have plenty of chance to talk over your difficulties with your parents.")

The sense of losing one's individuality in an anonymous organization is felt at all levels of seniority. An 18-year-old boy says that at school " you always feel restricted and socially unstable, obliged to follow all lines of authority with no expression of radical thought." The 18-year-old son of a joiner feels that " you have to try hard to preserve your individuality " ; the 17-year-old daughter of a tailor feels " disgusted that teachers do not take enough interest in the individual " ; the 16-year-old son of an insurance agent always feels " boxed, and one of a crowd " ; the 15-year-old son of a civil servant feels " part of a huge system."

The pressure of work and the aloofness and authoritarian attitudes of staff are mentioned as well as the impersonality of the organisation. "At school you always feel victimized," says a 17-year-old sixth former, son of a manufacturer's agent. It is, perhaps, more understandable that at 13 the daughter of an accountant should " always feel under their thumb."

A 15-year-old girl says she always feels as if she's in prison, " expected to obey the rules without questioning." The 16-year-old daughter of a principal viola player finds that " at school you always have plenty of chance to listen to your teacher's opinions " and " you always feel bored." (By contrast, at home " you always have plenty of chance to discuss things with your family," and " home helps you to develop your own interests.") At school, says the 15-year-old daughter of a wool-sorter, " you always feel that you are inferior to teachers—that is what they would like you to think." And at school you are expected to " work, work, work."

The 16-year-old son of a university lecturer echoes this sentiment : " At school you always have plenty of chance to slog, slog, slog, and to be treated like a kid who must not ask awkward questions." In his case, home brings no relief : " at home you always feel ignored." The club appears to be his only salvation : " at the club you always feel that you're enjoying yourself."

It may be the pressures and restrictions of the grammar school that account for the over-representation of grammar school pupils in youth clubs. This, of course, is only speculation ; but when we look at the pattern of expectation and satisfaction around home, school, club and work, it is possible to see how deprivations in one institution appear to find alleviation in another. Where school fails to satisfy, home, or club, may satisfy abundantly.

Tables 9 and 10 present in summary form the frequency of statements in the thirteen major need categories. Dominant needs or satisfactions, which are indicated by heavy type, are those categories which contain more than one-thirteenth of all statements.

TABLE 9

BOYS' STATEMENTS OF NEED AND SATISFACTION
(Percentages)

	Demand				Satisfaction			
	School (N : 336)	Club (405	Home (405)	Work (72)	School (N : 336)	Club (405)	Home (405)	Work (72)
Intellectual........	**23·8**	0·2	2·9	**13·9**	**30·9**	0·5	3·4	**13·9**
Physical	4·2	2·4	Nil	1·4	4·0	4·2	0·5	1·4
Manual	Nil	Nil	0·7	1·4	Nil	0·5	0·7	2·8
Social	2·9	**8·1**	2·7	Nil	1·8	5·2	1·0	1·4
Moral	**9·8**	4·7	**12·8**	**19·4**	5·0	5·4	**9·1**	**13·9**
Advancement......	**10·1**	0·6	5·2	**16·6**	**9·5**	0·3	3·9	**11·1**
Emotional security	**8·0**	**24·9**	**34·3**	6·9	*1·5*	**28·4**	**33·8**	**8·3**
Freedom	**22·0**	**13·5**	**22·5**	5·6	*3·6*	**10·4**	**13·3**	*2·8*
Friendship	3·9	**22·7**	1·0	7·0	2·6	**21·5**	1·2	**9·7**
Sense of competence	2·9	1·5	1·2	0·7	1·5	1·2	2·4	**8·3**
Support from adults	2·6	Nil	5·9	2·8	1·8	0·5	5·2	*Nil*
Identity with group	2·1	8·9	7·4	**15·3**	2·9	*3·7*	5·2	**11·1**
Sense of purpose ...	4·5	**10·3**	1·0	4·2	3·0	**10·1**	1·7	5·5
Negative statements					20·2	5·2	8·4	5·5

Legend :
Heavy type : major demand or satisfaction.
Italics indicate half or less of the demand.

TABLE 10

GIRLS' STATEMENTS OF NEED AND SATISFACTION
(Percentages)

	Demand				Satisfaction			
	School (N : 264)	Club (345)	Home (345)	Work (78)	School (N · 264)	Club (345)	Home (345)	Work (78)
Intellectual........	**22·7**	0·6	2·9	**14·2**	**21·9**	1·7	3·7	**15·4**
Physical	0·4	0·3	Nil	Nil	1·5	2·3	Nil	2.6
Manual	Nil	0·3	3·2	2·6	Nil	Nil	1·7	5·1
Social	7·2	**10·7**	2·0	3·6	*3·4*	**7·8**	1·5	1·3
Moral	**9·2**	2·1	**11·3**	**23·1**	7·2	2·1	6·4	**8·9**
Advancement......	**8·7**	0·6	3·7	**9·0**	*2·3*	0·3	2·3	**14·1**
Emotional security	**8·0**	**21·4**	**31·0**	**9·0**	5·3	**20·6**	**33·3**	**9·0**
Freedom	**22·3**	**13·3**	**23·9**	**7·7**	*8·3*	11·0	*11·3*	6·4
Friendship	5·7	**28·1**	3·8	**10·3**	6·8	**22·0**	2·1	**9·0**
Sense of competence	4·5	1·7	2·0	5·1	4·2	1·1	1·1	6·4
Support from adults	2·3	0·6	4·6	**7·7**	1·5	0·3	5·8	*Nil*
Identity with group	1·9	6·1	**8·1**	5·1	1·5	7·2	**9·3**	6·4
Sense of purpose ..	2·6	**11·0**	0·3	**9·0**	3·0	**11·9**	11·1	5·1
Negative statements					29·5	5·5	7·7	10·3

Legend :
Heavy type : major demand or satisfaction.
Italics indicate half or less of the demand.

VI.—CLUB MEMBERS AND NON-MEMBERS.

Youth club members are probably unrepresentative in various ways of the general population of young people between 14 and 20 years of age. Young people who are attending or have attended grammar schools appear to be members more often, in proportion to their numbers, than modern school pupils. From the evidence collected for the Crowther Report (1960), " It seems evident that, in respect of length of active membership, the boys from selective schools made an appreciably better showing throughout every type of youth organization than did others." It seems possible, on the face of it, that clubs attract the ' clubbable,' the more socially responsible and amenable of the teenage population.

In order to provide a rough check on the possible divergence of attitudes between the club members in this survey and other young people of similar age, schooling and occupation, a group of last-year secondary school pupils, and a group of young people at work, none of whom were club members, were invited to complete the appropriate sections of the questionnaire. The young workers were approached at recreational centres which were not youth organizations in any formal sense.

A last-year class in each of the three mixed secondary schools in the areas served by the clubs completed the questionnaire. Sixty-seven pupils (thirty-three girls and thirty-four boys) were not members of clubs. Their average age was 15 years. 85 per cent. came from the homes of manual workers, 15 per cent from white-collar homes.

These sixty-seven non-members were less negative in their attitude to school than club members of the same age who were attending the same type of school. Their demands of school were substantially the same, but their satisfaction with school was significantly different.

TABLE 11

STATEMENTS BY CLUB MEMBERS AND NON-MEMBERS ABOUT SCHOOL.

	Demand		Satisfaction		
	E	I	E+	E−	I
Non-members (N : 201)	44·7%	55·2%	39·8%	8·0%	52·2%
Club members attending modern schools (N : 210)	57·1	42·9	32·4	19·5	48·1

Chi square 12·0 d.f. 2 $P<0.01$

There was a similar difference in attitude to home. Non-members had the same expectations as members, but they made a significantly lower proportion of negative and critical statements.

Fifty young workers (twenty-five men and twenty-five women) between the ages of 16 and 20, who were not members of clubs, completed the questionnaire. Fifteen men were white-collar workers, ten were manual workers ; twenty-two women were in white-collar employment, three in manual. They did not differ from the club members who were in full-time employment in their attitude to work : 8·0 per cent. of their statements about work were negative, 7.3 per cent. of club-members' statements.

TABLE 12

STATEMENTS BY CLUB MEMBERS AND NON-MEMBERS ABOUT HOME.

	Demand		Satisfaction		
	E	I	E+	E−	I
Non-members (N : 201)	76·1%	23·9%	80·6%	1·5%	17·9%
All Club Members (N : 750)	77·2	22·8	72·3	9·7	18·0

Chi square 46·0 d.f.2 P<0·001

But they were far less negative in their attitudes to home. Only 2 per cent. of their statements were negative, compared with 8·2 per cent of statements made by club members of their age.

These results suggest that the youth club, far from attracting the less recalcitrant young people of the area, provide a refuge for those who feel more resentment toward their homes than other young people. This should not be interpreted as indicating that the homes of club members are generally ' unsatisfactory ' and severely deprived. Club members in general are highly appreciative of their homes, but to a rather less marked extent than the non-members who completed the questionnaire.

VII.—SELF CONCEPTS IN THE GRAMMAR SCHOOL.

The negative attitudes of grammar school pupils in general, and of senior pupils in particular, toward their schools, seemed to call for further invetsigation and, if possible, verification. If frustration of ' expressive ' needs is as widespread and severe as the survey seemed to indicate, it would be reasonable to expect that grammar school pupils, particularly in the upper school, would be characterized by generally negative self-concepts. The self-concepts of first formers and sixth formers were accordingly investigated in an apparently ' typical ' grammar school.

The school had not contributed any subjects to the original research population. It is a three-form entry mixed grammar school, recruiting pupils on the basis of selection tests at the age of 11. It is a well-established school with a good academic record.

Charles Cooley (1922) and G. H. Mead (1934) argued that a person's self-conception is derived from the way others define him and from the various positions he occupies in society. Mead stressed the importance of the ' generalized other,' the individual's conception of the organized social process of which he is part. If senior grammar school pupils in fact feel ' down-graded,' ' unimportant and hounded,' as some had alleged, we should expect this experience to be reflected in their self-conception.

Sarbin (1954) defined the self-concept as the sum of a person's self-perceptions, and argued that it could be discovered by asking subjects to indicate their qualities and characteristics on an adjective check-list. More recent work on the self-concept has included the subject's perceptions not only of himself but of others, 'A person's psychological field, his formulatiom of the world ' (Super, 1963). The W-A-Y test developed by Bugental and Zelen (1950), explores the self-concept in this wider sense.

Bugental and Zelen asked subjects to provide three answers to the question, " Who are you ? " They established seventeen response categories into which answers could be placed. Kuhn and McPortland (1954) used the same procedure but required twenty answers to their question. The author used the Kuhn and McPortland test with his grammar school subjects and classified the responses according to a slightly modified version of the response categories which have been used by other investigators. Subjects completed the test anonymously (although a small porportion, nevertheless, gave their names in answering the question).

The test was administered to 141 subjects, ninety in the first form (forty-two boys and forty-eight girls) and fifty-one (thirty-one boys and twenty girls) in the upper sixth. Seventy subjects came from the homes of manual workers, seventy-one from white-collar homes.

The seventeen response catergoies fall into a broad three-fold classification : (1) neutral statements, (2) affectively toned negative statements, and (3) affectively toned positive statements. Affectively toned statements may refer to self or to other people and external situations.

The self-concepts of sixth formers were far more negative than the self-concepts of first formers. The discrepancy was greater in the case of boys than of girls.

TABLE 13

BOYS' SELF CONCEPTS.
STATEMENTS BY FIRST AND SIXTH FORMERS.

	Negative	Positive	Neutral
First form (N : 840)	11·9%	28·2%	59·9%
Sixth form (N : 620)	30·2	28·0	41·8

Chi square 82·3 d.f.2 P<0·001

Sixth form boys made more affectively toned, and fewer neutral, statements than first formers ; the shift from neutral statements was entirely in the direction of negative comment. The same trend was apparent among the girls. There were no social class differences in the proportions of negative statements.

TABLE 14

GIRLS' SELF CONCEPTS.
STATEMENTS BY FIRST AND SIXTH FORMERS.

	Negative	Positive	Neutral
First form (N : 960)	14·3%	30·7%	55·0%
Sixth form (N : 400)	26·7	35·3	38·0

Chi square 42·6 d.f.2 P<0·001

Sixth form girls were not less negative than sixth form boys : their 26·7 per cent. negative statements are not significantly different from the 30·2 per cent. of the boys (CR 1·32, NS).

The first formers (they had spent two terms in the school) defined themselves in terms of their likes and enthusiasms, their membership of social groups and organizations. The sixth formers' self-concept was characterized by hatred of self and of others; by non-reference to membership groups. First formers tended to define themselves in terms of a geographical location, sixth formers did not. First formers saw themselves in physical terms; sixth formers (and girls even less than boys) seldom referred to their appearance. First formers referred to their family relationships more often than sixth formers; first form boys mentioned their age more often than sixth form boys. Sixth formers defined themselves in terms of socio-political affiliations (political parties, C.N.D. and the like) more often than first formers. There was no difference in the proportions of religious and metaphysical self-descriptions.

A half of the sixth form boys said they were drinkers and/or smokers; a third referred to their lack of money. A half of the sixth form boys referred to sex or members of the opposite sex: " I am sexually frustrated "; " I am in love "; " I like short, dark girls "; " I have had sex." Only two of the twenty sixth form girls made any reference whatsoever to the opposite sex (but a quarter of the girls said they were fond of children). One of the girls made three references to sex: " I am in love/I am in a personal dilemma concerning sex: is it always wrong before marriage? and is 18 too young to become engaged?/I am conscious of the fact that sometimes I do not love my boy friend, and more often he does not love me." A third of the boys at all ages and a sixth of the girls describe themselves as intelligent or highly intelligent. Tables 15 and 16 show the proportion of statements in the various response categories and the significance of the difference between first and sixth formers.

TABLE 15

STATEMENTS MADE IN THE W–A–Y TEST BY GRAMMAR SCHOOL BOYS.

Statement Categories	First Form (N : 840)	Sixth Form (N : 620)	Difference C.R.	P
A. NEUTRAL	%	%		
Name	2·5	0·3		
Pronoun	0·7	0·3		NS
Social-scientific	1·5	0·3	2·3	0·05
Metaphysical	0·1	0·3		NS
Sex	4·5	3·5		NS
Age	3·6	1·7	2·1	0·05
Occupation (future)	1·5	3·7	2·6	0·01
Family relationship	5·0	1·2	4·0	0·001
Group status/membership	10·1	4·3	4·1	0·001
Geographical-spatial	4·9	1·0	4·1	0·001
Socio-political	0·1	5·5	6·5	0·001
Nationality	1·5	2·5		NS
Religion	1·5	2·7		NS
Appearance	7·2	3·5	3·0	0·01
Neutral: undesignated	15·0	11·0	2·1	0·05
B. AFFECTIVELY TONED				
Negative	11·4	30·2	8·5	0·001
Negative-self	7·2	22·1		
Negative-nonself	4·2	8·1		
Positive	28·2	28·0		NS
Positive-self	11·4	17·4		
Positive-nonself	16·8	10·6		

TABLE 16

STATEMENTS MADE IN THE W–A–Y TEST BY GRAMMAR SCHOOL GIRLS.

Statement Categories	First Form (N : 960)	Sixth Form (N : 400)	Difference	
			C.R.	P
A. NEUTRAL	%	%		
Name	3·5	Nil		
Pronoun	0·5	0·8		NS
Social-scientific	1·5	2·6		NS
Metaphysical	0·2	0·8		NS
Sex	3·4	3·5		NS
Age	3·2	2·0		NS
Occupation (future)	2·2	4·0		NS
Family relationship	5·4	2·3	1·96	0·05
Group status/membership	9·7	6·3		NS
Geographical-spatial	3·0	0·8	2·3	0·05
Socio-political	0·2	3·4	6·5	0·001
Nationality	1·1	0·3		NS
Religion	2·0	2·3		NS
Appearance	5·1	1·5	3·0	0·01
Neutral : undesignated	14·0	7·4	6·3	0·001
B. AFFECTIVELY TONED				
Negative	14·3	26·7	5·2	0·001
Negative-self	8·0	24·0		
Negative-nonself	6·3	2·7		
Positive	30·7	35·3	2·4	0·05
Positive-self	7·3	15·5		
Positive-nonself	23·4	19·8		

First form pupils made far more (neutral statements which were impossible to classify in the established categories (see ' undesignated : neutral ' responses). The self percepts of sixth formers were far more homogeneous than the self percepts of first formers. The latter tended more often to describe themselves in unclassifiable ways, such as " I am a Leo subject," " I am a vegetarian," " I am a person without an appendix." (Medical histories figured in their self-conceptions, almost never in sixth formers : " I am a dentist hater," " I am always having colds," " I am a doctor hater," " I am on a hospital waiting list."

Religion was not prominent at any age with either sex ; but, in one or two cases, it dominated the self-image. A 12-year-old girl described herself as " one of God's children/a sinner in God's name/a chosen person of God/a member of the congregation of the Central Methodist Church/a member of the Sunday school." (She is also a girl guide and hopes to go to a university to study Sociology.)

Negative self-references in the first form refer more often than in the upper sixth to personal appearance and performance in school subjects and other school activities. " I am lumpy/I am plain " ; " I am not a very good swimmer/not a very brave person/frightened of the dentist/not very good at making friends quickly/not very good at making conversation." Social uncertainty is apparent here, but is still more prominent six years later. A first form boy sees himself as " a show-off/weak willed/always in trouble/rotten at French/hopeless at cross-country running/a lazy person really/dim compared with my friend/untidy/fed up with my sister." But much more commonly at this age the self-conception is predominantly positive : " I am very lucky to have such a good home/I am much indebted to my grandparents for what they have done for me/I am very fond of the country-side/I am a lover of cats/I am a lover of snow, ice and camping out."

There is little joy in the upper sixth. Only one predominantly positive self-image emerged, with positive statements outnumbering negative. This was a girl who saw herself as " a highly intelligent person/interesting/a poet/unique/ in love with life/glad I am myself/self-confident/a person who usually succeeds." But even she concedes : " I am often rather snobbish/I am inclined to attempt too much at once."

The following excerpts illustrate the negative content of ten sixth-form boys' self-conceptions :

(1) " I am unsure of my abilities/afraid of failure/a person who needs many friends/concerned about my A-levels/worried about my girl/selfish." On the other hand, he thinks he is " lucky/tall, and stronger than most/fond of drink and women."

(2) " I am greedy/self-centred/weak willed/vain/unable to accept the world/shoddy in appearance/full of self pity/ashamed of myself." But he also sees himself as " intelligent/kind/generous/in love."

(3) " I am against old people, coloured people, Jews/I find most adults annoying/I hate my father/I do not think young people owe their parents anything/I hope never to live to be an old man/I hate zoos and circuses." But he " drinks and smokes 90 cigarettes a week," and " has had sex ".

(4) " I am dissatisfied with my present social status/worried about my future/disliked by a large portion of the lower school/feel lonely at school/am not skilled at anything/not a member of a club of any sort/not attractive in appearance/looking forward to leaving school." He is a Conservative in politics and has " several high but improbable ambitions."

(5) " I am against coloured immigrants/I am anti-Yank/an agnostic." (But he, too, is a drinker and a smoker and attracted by women.).

(6) " I am a coward/lazy/quick tempered/a hater of cats."

(7) " I am afraid of being physically crippled " (but he is " fond of short, dark girls ").

(8) " I am frightened by loneliness/I am even more frightened by large social gatherings/I am disgusted with other people's self-confidence/pessimistic by nature/tempted and lack will power/a bad loser/in many ways a hyopcrite/undecided concerning hanging and euthanasia/a church member who feels wholly inadequate."

(9) " I am in favour of the colour bar/a person who would not stand out in a crowd/rather lazy/a person who looks down on Communists."

(10) " I am a segregationist/I am disappointed with many aspects of my life." But he's a drinker and a smoker and (perhaps for this reason) " a person not in the best of health."

The girls were similar in their self-appraisal. The following is illustrative material :

(1) " I am easily depressed/in a rut/afraid of death/mixed up as regards my religious beliefs/unable to relax/dissatisfied with my present station in life." On the other hand, she says she is intelligent and hopes to go to college.

(2) " I am sometimes afraid/uncertain about the future/happiest when I forget about my present life/able to forget by singing or playing a musical instrument/shy till I get to know people/hoping to pass my exams., but don't think I will/afraid of lots of things, like the dark/happy to sit in front of a roaring fire rather than go out."

(3) " I am anxious over exams./not glamorous/not very kind to human beings, but kind to animals."

(4) " I am hopeless at making decisions/not very good at ordering people about, because they take no notice of me/I am dithery." (But she likes peanuts, liquorice, and the sea.)

(5) " I am often very depressed/conscious of a desire to be popular/ hoping to go to the university, but scared at the prospect."

(6) " I am uncertain about the future/afraid of death/extremely lazy/ extremely plain." (But she is " a good cook " and " very partial to Vodka.")

(7) " I am a pessimist/an introvert, unfortunately/a basically shy person/ a highly strung person/a very moody person." She intends to be a teacher.

(8) " I am selfish/a snob/too fond of the limelight/lazy/inconsistent." She intends to be a teacher, too.

VIII.—CONCLUSION.

This inquiry suggests that home and work may be meeting the needs and expectations of young people to a greater extent than is often supposed ; and that the grammar school, particularly in its higher reaches, may often be failing to promote the self-confident and assured maturity which is often supposed to develop from sixth form experience. There were no indications in this inquiry that there is, in general, a great gulf between teenage children and their parents ; that they are unable to communicate ; that children throughout their teens do not place high value on their parents' support and approval. All the evidence is to the contrary. Between 14 and 20, the young people attending six northern youth clubs—and other groups who were not members of clubs—in general appeared to appreciate their homes and to find in them ease, relaxation and re-assurance.

Youth clubs do not appear to be the arenas of strenuous self-improvement which youth leaders would like them to be ; but they offer a sense of release and relaxation which other institutions provide less abundantly. They are also valued for the opportunity they afford for developing social skills and extending one's social horizons.

Any interpretation of grammar-school pupils' negative attitudes to school can only be speculative. The grammar school measures up to the considerable demands for intellectual development and personal advancement. But it is seen as a largely restrictive institution, and many pupils appear to find staff-student relationships remote and chilling even in the sixth form.

The negative self-concepts of sixth formers are in line with this picture. The developed critical powers of senior pupils might be held to account for their comparatively unfavourable view of themselves and their schools. But they do not view their homes and clubs less favourably than younger or less intelligent people. And the hostile comments which appeared in the W-A-Y test were often more suggestive of blinkered, stereotyped thinking (for example with regard to coloured immigrants and other social problems) than of an enlightened, radical and informed re-appraisal of life and society.

The sixth form probably tends to select negative and introverted personalities. Their negative and introverted characteristics are perhaps reinforced by their experiences. Negative self-concepts may be a necessary price of high academic attainment under highly competitive conditions. And it may be the case that some discrepancy between expressive needs and expressive satisfactions constitutes a creative tension. But there is probably a point beyond which this tension ceases to be creative and becomes demoralizing.

No attempt was made in this inquiry to relate expressive need-frustrations to academic performance. There was no indication in these data that those who were most critical of school came from the 'poorer' schools, with generally lower attainment; there were, indeed, some signs that the reverse was true. Numbers were not large enough to make inter-schools comparisons possible but there appeared to be a tendency for pupils from large grammar schools of high academic distinction, recruiting their pupils in the main from higher professional families (there were four such schools contributing subjects) to be particularly negative in their attitudes. The inquiry did not begin as a research into the personal consequences of an advanced academic education. But the material which emerged suggests that grammar schools which differ in their size, social composition and academic distinction might be usefully compared for the attitudes and self-conceptions which they induce in their pupils.

ACKNOWLEDGEMENTS.—I am indebted to Mr. G. H. Johnson, B.Sc., and Mr. G. B. Miles for their invaluable help in the administration of this inquiry.

IX.—REFERENCES.

BUGENTAL, JAMES, F. T., and ZELEN, SEYMOUR L. (1950). Investigations into the self concept, I.—The W-A-Y technique. *J. Personality*, 18, 483-498.

COOLEY, CHARLES HORTON (1922). *Human Nature and the Social Order*. Glecoe, Ill.,: The Free Press.

KUHN, H. MANFORD and MCPORTLAND, THOMAS S. (1954). An empirical investigation of self attitudes. *Amer. Sociol. Rev.*, 19, 68-78.

MEAD, G. H. (1934). *Mind, Self and Society*. Chicago University Press.

SARBIN, THEODORE R. (1954). Role theory. In LINDZEY, G., *Handbook of Social Psychology*, Vol. I. Massachusetts: Addison-Wesley.

SUPER, DONALD E. (1963). *Career Development: Self Concept Theory*. New York: College Entrance Examination Board.

Fifteen to Eighteen (1960). H.M.S.O.

(*Manuscript received 13th May, 1965*)

DIFFERENCES IN TEST SCORE AND IN THE GAINING OF SELECTIVE PLACES FOR SCOTTISH CHILDREN AND THOSE IN ENGLAND AND WALES

By J. W. B. DOUGLAS and J. M. ROSS
(*Medical Research Council Unit, London School of Economics*)
AND
S. M. M. MAXWELL and D. A. WALKER
(*The Scottish Council for Research in Education*)

SUMMARY. The National Survey is a longitudinal study of 5,000 boys and girls born in 1946. In this paper the test scores and allocation to selective courses of children living in Scotland are compared with those in England and Wales. The Scottish pupils were superior to the English and Welsh in the tests of sentence completion and reading at 8 years, in arithmetic at 11 years and in mathematics at 15 years, whereas the English and Welsh were superior in tests of non-verbal intelligence at all three ages and in vocabulary and verbal intelligence at 11 years.

It is suggested that the growing apart of the social classes in test results between 8 and 11 in England and Wales, which did not occur in Scotland, may arise from the pressure associated with secondary selection at 11+. Social inequalities in selection were found both North and South of the Border, being slightly greater in the South. The more generous provision of places in Scotland is demonstrated by the fact that the Scottish manual working class boy or girl at each level of ability was approximately as likely to embark on a selective course as a middle class pupil in the South.

I.—INTRODUCTION.

FOR the past nineteen years a continuing study has been made of the growth and development of 5,362 children born during the first week of March, 1946. These children come from all parts of England, Wales and Scotland, but the first report of their educational progress, published in 1964 as " The Home and the School," was confined to the English and Welsh children. This was because selection for secondary education takes place a year later in Scotland than in England and Wales, a greater proportion of the Scottish children follow academic courses and the selection procedure and tests used in the two countries are not directly comparable. A number of requests for information about the progress of the Scottish children in the primary schools have been made, and this brief account seeks to fill in some of the gaps and compare the test results and the social class differences in secondary selection in the two countries.

II.—SCHOOLS AND COURSES FOUND IN THE TWO COUNTRIES.

There are considerable differences in primary education North and South of the Border. The Scottish primary course covers seven years instead of six, and, while there are separate infant and junior departments, it is rare for them to be housed in different schools. Transfer from the infant department to the next class may well take place in mid-session and need not be deferred until the beginning of the Autumn term. The secondary school organisation is also different from that in England and Wales. In Scotland there are no technical schools, and in 1958, when the National Survey children reached 12 years, they were allocated to either senior secondary (i.e., five-year) or junior secondary (i.e., three-year) courses. The senior secondary courses correspond roughly to those in the English grammar or technical schools and the junior secondary courses approximate to those in secondary modern schools. Both types of course are frequently provided in the same school. For convenience, courses of the senior secondary, grammar or technical type and those provided in grant-aided or direct grant schools are described as selective courses or schools. In 1958, when the Survey children were twelve, 38 per cent.* of the Scottish members were in selective courses or schools as compared with 30 per cent.* of the English and Welsh children in the Survey.

The much larger proportion of places available for pupils desiring selective secondary courses is almost certainly one factor easing pressure in the top classes of Scottish primary schools. Any parent dissatisfied with the education authority's allocation has the right to appeal to the Secretary of State: in 1958 the number of appeals was seventy-four in an age-group of 90,000. There is no doubt, however, that pressure exists and is more intense in some schools and some areas than in others, but it would seem to be less intense, on the average, in Scotland than in England and Wales. Finally, there is the intangible tradition, which the Scots like to think persists, that the way is open for any able pupil, whatever his social class or economic background, to reach the university.

The following analysis is divided into two parts: the first is concerned with the test results of the Scottish children on the one hand, and the English and Welsh on the other, and the second part takes a closer look at the question of the social class differences found in allocation to selective courses.

III.—THE TEST SCORE DIFFERENCES.

(a) *The Tests and the Sample.*

The members of the National Survey of Health and Development sat three test batteries at 8, 11 and 15 years. Each test was standardised on all the boys and girls who took the test, after allowing for the fact that only one in four of those from the manual working class were included in this follow-up study. The mean for each of the tests is 50 and the standard deviation 10. At 8 years, the tests were of picture intelligence, sentence completion, vocabulary and reading: at 11 years they were of general ability, arithmetic, vocabulary and reading, while at 15 years they were of intelligence (both non-verbal and verbal), mathematics and reading.†

For each of the three years a mean T score is also available. This measure includes both ability and attainment. The mean T scores were obtained by adding the four standardised results at each age and then re-standardising this aggregate score on the group that sat all three sets of tests (the population estimate number for those with complete educational information is 8,708: the actual number is 3,626). The mean T score for each year is 50 and the standard deviation 10.‡

---

* These figures, which have been corrected for sampling, include the children at independent and private schools. The correction was needed because the sample included all the births to non-manual workers' and agricultural workers' wives, but only one-quarter of the births to wives of other manual workers occurring during the survey week. These figures corrected for sampling are known as 'population estimates.'

† For a full description of the Survey and the tests used at 8 and 11, see "The Home and the School," Appendix I, by D. A. Pidgeon.

‡ These aggregate T scores are different from those used in "The Home and the School" which, not being re-standardised, had standard deviations of 8.5 and 8 9 respectively, at 8 and 11 years.

It is the group of boys and girls with full test score information that is used in this analysis. The means of the individual tests for this group are slightly above 50, with one exception—this is because the pupils who did not complete all three batteries tended to do less well in the tests they did sit. The intelligence test at 11 years can be divided into two halves—a non-verbal and a verbal, but it is the total only that is used as one of the four tests in the calculation of the mean T scores for that age. The intelligence tests at 15 years—a non-verbal and a verbal—have equal weight and form two out of the four tests used in calculating the mean T scores for that age.

19·5 per cent.* of the sample with full test information in Scotland come from the middle class, as compared with 24·5 per cent.* in England and Wales. It seems inappropriate, when comparing the two countries, to standardise for social class, though the social class disparity should, of course, be borne in mind when any results are compared. There appear to be slightly more boys in the Scottish Survey—53·6 per cent.* against 50·6 per cent.* in England and Wales, though this difference is not statistically significant.

---

* Figures adjusted for sampling (see note p. 151).

TABLE 1

A COMPARISON OF TEST SCORES AT THREE SEPARATE AGES OF THOSE CHILDREN LIVING IN SCOTLAND AND IN ENGLAND AND WALES, DIVIDED BY SEX AND SOCIAL CLASS—POPULATION ESTIMATES.

	SCOTLAND					ENGLAND AND WALES				
	Both Sexes		Both Social Classes		Both Social Classes and Both Sexes	Both Sexes		Both Social Classes		Both Social Classes and Both Sexes
AGE OF TESTING	Social Middle	Class M–W	Boys	Girls		Social Middle	Class M–W	Boys	Girls	
(*a*) 8 years.										
Picture Intelligence	53·6	47·5	48·8	48·5	48·7 †	53·6	49·5	50·5	50·5	50·5 †
Sentence Completion ..	57·2	50·7	52·2	51·7	52·0 †	54·7	48·6	49·4	50·7	50·1 †
Vocabulary	57·1	48·5	50·2	50·1	50·2	55·1	48·7	50·7	49·8	50·2
Reading	57·6	51·7	52·9	52·8	52·9 †	54·5	48·4	49·2	50·6	49·9 †
Mean T.	57·0	49·2	50·8	50·5	50·7 *	54·8	48·2	49·5	50·1	49·8 *
(*b*) 11 years.										
Non-verbal Intelligence ..	52·0	46·4	47·4	47·5	47·5 †	55·2	48·8	50·2	50·5	50·4 †
Verbal Intelligence	53·6	47·3	47·6	49·6	48·6 †	55·3	48·7	49·2	51·5	50·3 †
Intelligence Total	53·2	46·9	47·5	48·8	48·1 †	55·6	48·8	49·7	51·2	50·4 †
Arithmetic	56·6	51·7	52·3	53·0	52·6 †	54·6	48·3	49·6	50·1	49·9 †
Vocabulary	55·3	46·8	48·8	48·1	48·5 †	55·6	48·5	50·8	49·7	50·3 †
Reading	55·5	49·5	51·2	50·0	50·7	55·0	48·4	49·7	50·4	50·1
Mean T.	55·4	48·4	49·7	49·7	49·7	55·6	48·2	49·8	50·2	50·0
(*c*) 15 years.										
Non-verbal Intelligence ..	51·2	48·0	49·4	47·6	48·6 †	53·8	50·0	51·7	50·1	50·9 †
Verbal Intelligence	54·4	49·0	49·6	50·7	50·1	54·4	48·8	49·7	50·7	50·2
Intelligence Total	53·0	48·2	49·3	48·8	49·1	54·5	49·1	50·6	50·3	50·4
Mathematics	56·3	50·1	52·6	49·9	51·3 †	55·1	48·2	51·4	48·4	49·9 †
Reading	57·0	48·6	50·7	49·6	50·2	55·5	49·1	51·1	50·3	50·7
Mean T.	55·2	48·2	50·2	48·9	49·6	55·2	48·4	50·7	49·4	50·0
No. (population estimate)	213	882	587	508	1095	1862	5751	3850	3763	7613

* P<0.05. † P<0·01.

(*b*) *Test Results for the two Countries Compared.*

Eight-year Test Summary. The Scottish children's average test scores were higher for both reading and sentence completion : the differences of 3·0 and 1·9 for the two countries taken as a whole, are highly significant ($P<0{\cdot}01$). The Scottish children's average result for picture intelligence was, however, lower and this difference of 1·8 in favour of the English andWelsh children is also highly significant ($P<0{\cdot}01$). Over the four tests the Scottish groups meanT score was 0·9 points higher than that for England and Wales ($P<0{\cdot}05$).

Eleven-year Test Summary. The English and Welsh 11-year-olds had a very slightly higher mean T score than the Scots, but the difference of 0·3 points is not statistically significant. The results of three of the individual tests did show, however, differences between the two countries that are highly significant ; the English and Welsh children did better in the general ability test (2·3 points), the discrepancy being larger for the non-verbal than the verbal part of this test, and they also made an average score in the vocabulary test that was 1·8 points above that of the Scottish children, whereas the latter made higher scores in the arithmetic test, being 2·7 points above the English and Welsh.

Fifteen-year Test Summary. The mean T scores of 15-year-old children in the two countries did not differ significantly, though there were two highly significant differences in the individual test scores, which, being in the opposite direction, cancelled each other out in the aggregate score. The first was for the non-verbal intelligence test, where the English and Welsh children were 2·3 points above the Scottish and the second was the mathematics test, where the Scottish children were superior by 1·4 points.

The Scottish children in this national study were, as a whole, superior to the English and Welsh in attainment tests given in the early primary school years ; they made particularly high scores in the reading and sentence completion tests. By 11 years the English and Welsh children had caught up with the Scottish children in reading, and surpassed them in vocabulary, but the Scots had the higher average arithmetic result, and retained this advantage in the mathematics tests given at 15.

The English and Welsh children showed themselves to be superior in tests of non-verbal intelligence at all three ages.

The significantly higher mean T score of the Scottish children at 8 years, followed by a relative fall in the next 7 years, is difficult to interpret because, of course, the tests used varied from year to year, not only in subject matter, but also in type.* It seems, however, that the Scottish children, who at 8 years, were superior as a whole to the English and Welsh, lost ground to the extent of 1·1 points—or just over one-tenth of a S.D. during the seven subsequent years' schooling, while the English and Welsh improved their overall score very slightly. In other words the early advantage of the Scots of 0·9 points of mean T score at 8 years had vanished at the age of 15 years.

(*c*) *Differences in the Test Performance of Boys and Girls.*

The English and Welsh girls made higher scores than the boys in all the tests given at 8 and 11 years, with the exception of the two vocabulary tests. The Scottish boys did as well as the girls at both these ages. In the fifteen-year tests the girls in both countries fell behind the boys in non-verbal ability, reading and mathematics. In the verbal intelligence test, however, the girls on the average made higher scores than the boys. On the aggregate of all tests the average 15-year-old girl was 1·3 points below the average 15-year-old boy: exactly the same difference was found North and South of the Border.

---

* There are two exceptions to this—the reading and vocabulary tests at 8 and 11 were the same.

(*d*) *Social Class Differences in Test Score.*

In examining the changes between 8 and 11 in the performance of the two social class groups in the two countries, the investigator is faced with some difficult statistical questions. The obvious one is that of regression to the mean; this would tend to narrow the gap between two groups which at first selection were, respectively, above and below the general average. The problems have not yet been resolved and the following account is, therefore, descriptive, and the relative differences have not been tested for statistical significance. A further difficulty in interpreting the changes in performance over time is that, with two exceptions, the tests used differed at each age and, even with the mechanical reading and vocabulary tests, which were the same at 8 and 11 years, it is by no means certain that the assessments are truly comparable.

In England and Wales, the middle and manual working class moved further apart between 8 and 11 years, in their performance in both the vocabulary and the reading test. In vocabulary, the middle class children were 0·7 points of score further away from the manual working class children at 11 years than at 8; in the reading test they were 0·6 points further away. The average mean T score of these two social groups moved apart by 0·8 points between 8 and 11 years.

In contrast, in Scotland, the gap between the social classes for vocabulary and reading showed little change between 8 and 11 years, the middle class being 8·6 points ahead of the manual working class for vocabulary at 8 years and 8·5 points ahead at 11. In reading, the middle class children in Scotland were 5·9 points ahead of the manual working class at 8 and 6·0 points ahead at 11. When the mean of all test scores at the two ages is taken, the middle and manual working classes in Scotland appear to converge, being 0·8 nearer together at 11 than at 8. This is largely owing to the replacement of the sentence completion test, which showed wide social class differences at 8, by an arithmetic test at 11, which showed much smaller differences.

This differing social pattern between Scotland and England and Wales may, perhaps, be explained, in part, in terms of the pressures exerted by the English system of selection examinations at 11+ and partly, perhaps, as showing the effect of practice on test performance. Another interpretation is suggested by the observation that at 8 years the middle class children in Scotland did considerably better than the middle class children in England and Wales in both the reading and the vocabulary tests, though they did not maintain their advantage when the same tests were repeated at 11 years. It may be that the pressures on middle class children in Scotland are concentrated in the early years at primary school and in England and Wales at the point of secondary selection.

Between 11 and 15 years the social class differences in the average mean T scores narrowed slightly in both countries. This is largely owing to changes in performance in the verbal and non-verbal tests. The social classes in Scotland diverged substantially in both the reading and the mathematics tests, whereas in England and Wales they converged slightly in reading and diverged slightly in mathematics.

IV.—SOCIAL CLASS DIFFERENCES IN ALLOCATION TO SECONDARY COURSES

The children described in this section are all those for whom there is both 11-year test and school allocation information. The sample is, therefore, slightly larger than that just described, which was restricted to those who had completed all three of the test batteries. The population estimate figure for the following discussion is 8,895 (for the previous section it was 187 fewer). It should also be mentioned that the S.D. of the mean T scores given in Table 2 is 8·9; this allows the figures to be compared directly with those appearing in "The Home and the School."*

The numbers of children allocated to selective courses are shown in Table 2. A selective course in Scotland is defined as a senior secondary or five-year course, as described in the second section of this article, or a course in a grant-aided school. In England and Wales, all those pupils who, at 11 years, were at grammar, direct grant, technical, independent or private schools, or who were in grammar or technical streams of bilateral, multilateral or comprehensive schools are regarded as receiving selective education.

TABLE 2

PROPORTIONS IN SELECTIVE SCHOOLS OR SELECTIVE COURSES

(Population Estimates)

Mean T-score at 11	SCOTLAND					ENGLAND AND WALES				
	Middle Class		Manual Working Class			Middle Class		Manual Working Class		
	Total No.	% on Selective courses (a)	Total No.	% on Selective courses (b)	Ratio a/b	Total No.	% on Selective courses (a)	Total No.	% on Selective courses (b)	Ratio a/b
60 & over	47	97·9	28	100·0	0·98	602	96·0	415	94·2	1·02
55-60	69	92·8	139	80·6	1·15	668	70·4	979	51·2	1·38
52-54	59	69·5	132	34·8	2·00	285	36·1	691	18·8	1·92
46-51	53	28·3	281	16·0	1·77	403	15·6	1374	6·0	2·60
0–45	42	0·0	230	7·8	0·00	332	5·4	2066	1·4	3·86
Total ..	270	61·5	810	30·7	2·00	2290	53·8	5525	20·5	2·62

In interpreting the table, the following points should be born in mind. The allocation to courses in either country was not based on the T score, but was determined by a procedure varying from area to area. The correlation between the T scores and the measures employed is probably high, but not nearly perfect. This partly accounts for a number of children with relatively low T scores being allocated to five-year courses or selective schools, and for some with relatively high T scores not being allocated to these courses. Secondly, there is the operation of parental choice. In the immediate post-war years, there were many children in Scotland who were deemed fit to enter five-year courses, but whose parents chose otherwise. That position altered considerably in subsequent years. The data in Table 2 give no indication of the extent to which this factor operated in 1958, but a search of the Survey records has produced only two instances of a place in a five-year course being refused.

---

* The S.Ds. of all the mean T scores were adjusted to 10 at a later date.

The first and most noticeable feature of the table is the much higher proportion of children in Scotland in each social class who were allocated to the selective type of course—38 per cent. against 30 per cent. This is not a question of more accommodation or of more places for this type of course being available in Scotland. The borderline for admission was fixed in most areas of Scotland by applying the results of research, especially that of McClelland (1942). At the time when the findings embodied in his report were being transformed into administrative principles, the proportion of an age-group estimated to be fit to profit from the five-year course was 25 per cent. Within a few years, the criterion on which the estimate was based was changed by new regulations affecting five-year courses and the types of certificate awarded at the end of them, and the percentage of an age-group estimated to be fit to profit had to be increased. Recent changes have again affected the criterion and the position is once more under review.

The second feature of Table 2 is the relation between social class and entry to a selective course or school. Here the comparison between the two countries is complicated by the independent schools which provide a substantial proportion of the selective places in England and Wales (5 per cent.) and a very small proportion (less than 1 per cent.) in Scotland. As they are almost exclusively middle-class schools, it would be expected that there would be much greater social inequality South of the Border. Above a moderately high standard of ability represented by a T score of 60, i.e., approximately one standard deviation above the average, social class made little or no difference in either country. In the next band, where the T score range is from 55 to 60, the chances of the middle class child, compared with those of the manual working class child, were in the ratio of 1·15 for Scotland and 1·38 for England and Wales. This is nearing the border zone when it was, at that time, doubtful whether the child would make a success of the selective secondary school course. The greater ratio in England and Wales may represent the relatively greater willingness of middle class parents in the South to take the risk of seeking places in independent schools, for children who had failed to gain grammar or technical school places. However, the relative nature of this comparison should be underlined, for the proportion of manual working class children at each level of ability (except one) in selective courses in Scotland exceeds the proportion of *middle* class children of similar ability in selective courses in England and Wales. On this hypothesis one would expect an increasing deviation in the ratios of the two countries, as the T score fell, and this is so in the lowest bands.

A further possible explanation of the differences between the two social classes is that the measures used by authorities in allocating pupils to selective courses contained some element, absent from the T scores, and positively correlated with social class. One such element, in Scotland, might be teachers' estimates of attainment in English and Arithmetic which are widely used as part of the selection battery. In England and Wales the selection procedure used in 1957 differed greatly from one area of the country to another, but many areas took account of teachers' assessments of the children's ability.

To summarise the social inequalities in selection : the position in the two countries was similar except that at almost every level of measured ability and in both social classes the Scottish child had a higher chance of being found on a selective course—indeed, at each level of ability, the manual working class child in Scotland was approximately as likely to enter a selective course as the middle class child in England and Wales. The rather greater social inequalities recorded in England and Wales are to be expected owing to the large part played by independent schools in providing selective education in that country.

ACKNOWLEDGEMENTS.—The National Survey of Health and Development was sponsored by a joint committee of the Institute of Child Health, the Society of Medical Officers of Health and the Population Investigation Committee. It is now being directed from the Medical Research Council Unit at the London School of Economics. In its early years the inquiry was financed by the Nuffield Foundation, and grants for continuing it through the school years have come from many sources, including the Ford Foundation and the Population Council, Inc.

We have received much assistance from the National Foundation for Educational Research in England and Wales—they supplied us with all the school tests and both checked the markings and standardised the scores. We wish to thank them for all their help.

A National Longitudinal Study such as this is possible only because of the generous co-operation we have had from many people and agencies: here we would especially like to thank the directors of education, teachers and health visitors in all parts of Great Britain, who have given us information about the boys and girls included in this Survey. We should also like to thank the mothers who willingly answered numerous detailed questions, and the boys and girls themselves.

V.—REFERENCES.

J. W. B. DOUGLAS (1964). *The Home and the School.* London: MacGibbon and Kee.

W. MCCLELLAND (1942). *Selection for Secondary Education.* Publications of the Scottish Council for Research in Education XIX. London: University of London Press.

(*Manuscript received* 15*th July,* 1965)

THE RELATION OF TEACHING ABILITY TO MEASURES OF PERSONALITY

By K. B. START

(Research Fellow, School of Education, University of Manchester)

Summary. Thirty-five members of staff of the mixed secondary modern school completed the 16 P.F. Test and these scores were compared with those of teaching ability as rated by the head teacher. It appeared that the best teachers differed from the rest on factors A –, B+, E+, L –, M+, Q1 – and Q3 –. As teaching ability was linked with sex, correlations partialling out sex indicated that factors L – and M+ were significantly (·05) correlated with rated teaching ability and that H+ marginally failed to reach this level of significance.

I.—Introduction.

The inability to agree as to which measurable traits can be used to differentiate between effective and ineffective teachers is a major handicap in research in this particular area of educational psychology. The crux of the problem appears to lie in the wide interpretation of 'success' as a teacher and the multifarious, mainly subjective, criteria which are used individually in some cases but more commonly in varying combinations to decide whether a teacher is efficient or not.[1] Studies relying on the subjective opinion of a variety of judges—often head teachers—are bedevilled by this problem of multiple and different criteria. In the present study an attempt was made to circumvent the difficulty by having one rater and hence one set of criteria to contend with. The intrinsic errors of bias which exist in these circumstances are obvious but it was hoped that it would be possible to establish objectively some of the criteria used in assessing teacher ability in a particular school situation as distinct from teaching practice assessment.

II.—Method.

The staff of a mixed, secondary modern school in Lancashire agreed to complete the 16 P.F. Test, the Manchester Educational Attitude Scale[4] and a brief biographical questionnaire. The Head Teacher of the school was asked to rate the Staff of thirty-nine on nine criteria, one of which was general teaching ability, so that they conformed to the following distribution :

Among the best members of my staff	Rather above the average members of my staff	About average among members of my staff	Rather below the average members of my staff	Among the weakest member of my staff
top 5	Av.+ 8	Av. 14	Av.– 8	bottom 5

Unsealed but numbered envelopes were provided each enclosing single copies of the 16 P.F. Test, the Manchester Educational Attitude Scale and the short biographical questionnaire. All three tests were stapled together. The Head Teacher issued these envelopes to the Staff and provided the staff ratings identified only by the envelope number. The envelopes were returned sealed to the Head Teacher who could not, therefore, learn any opinion of a staff member. The author related envelope contents with the Head Teacher's rating by means of the envelope number and in this way the anonymity of individual members of staff was ensured.

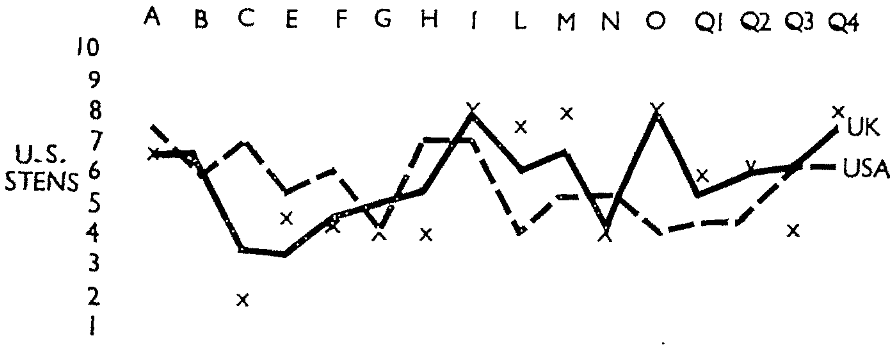
The Head Teacher's ratings were scored 5, 4, 3, 2, 1 in descending order for each of the nine criteria, viz., teaching ability, personal relations with staff colleagues, class control, personal relationship with parents, administrative ability, personal relations with children, extra-mural activity and personal relations with head teacher. In examining the inter-correlations between the ratings on the different criteria an obvious halo effect was seen as 29 of the 36 correlations were in excess of ·5. The unweighted American sten scores of the ·16 P.F. Test were calculated. Since the professional estimate of teaching ability and the personality of the teacher were the major interest only these two aspects of the study are discussed in this paper, except where specific reference to some other item throws light on the 16 P.F. data.

III.—RESULTS.

Thirty-five of the thirty-nine members of staff responded to the enquiry. The ratings of the four staff members who did not co-operate placed one in each of the top, above average, average and bottom groups. The distribution of the ratings did not closely follow the suggested pattern as it proved difficult for the Head Teacher to make sufficiently discrete comparisons and resulted in a distribution, in descending order of teaching ability, of 7, 6, 11, 7, 4. The data were analysed by teaching ability, by sex, by sex and ability, and by teaching ability with sex partialled out.

First Order Correlations.

The Mean scores for the total group on each of the 16 P.F. factors appear in Fig. 1, with those of fifty-nine elementary and junior high school teachers from the U.S.A.



x—Means based on 1221 UK Subjects, by permission of Professor FW Warburton, Manchester University

FIGURE 1

Without access to the raw data on the American sample, tests of significance between the two groups of teachers are impossible. It would seem that differences on Factors C (Emotional Security), E (Dominance), L (Protension) and O (Guilt Proneness) are interestingly large, but these might reflect the mean scores from the larger British sample, which differ markedly from the American scores on these traits.

The data were then separately analysed by ability groups and the profiles of good(5), average (3) and weak (1) teachers appear as in Fig. 2.

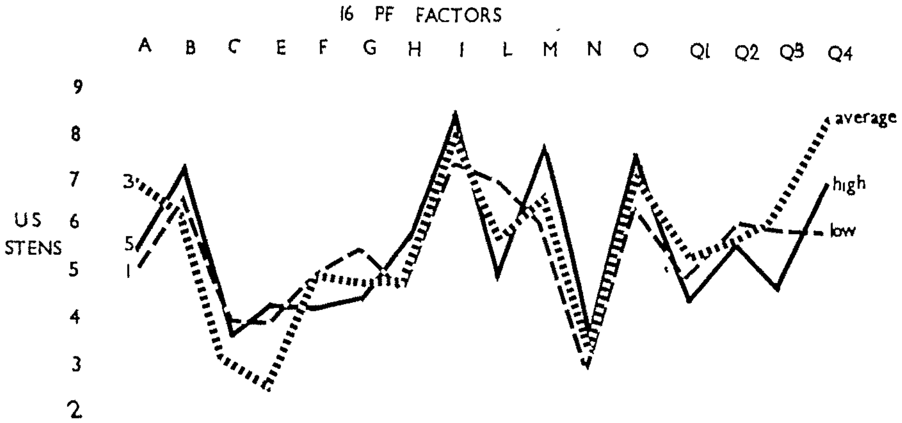


FIGURE 2

Inter-group comparison is hazardous because of the low numbers involved, but routine F, t and C.R. tests were carried out on groups and likely combinations of groups. The data were examined to see if good teachers could be differentiated from the rest and whether such a division could be seen between poor teachers and their colleagues. The latter comparison was less profitable because group 1 (weakest teachers) tended to have a high variance. Differences seemed to exist between poor teachers and their colleagues on factors C, G, I and Q4, but the C.Rs. were 1·40, 1·30, 0·80 and 1·58, respectively. A more fruitful approach derived from comparing the best teachers with the rest of the staff. The significant differences which were observed are tabulated in Table 1.

TABLE 1

SIGNIFICANT DIFFERENCES OF TEACHING ABILITY GROUP MEANS ON THE 16 P.F. TESTS.

Factor	Groups Compared	Tests	Sig. Level	Direction of Group 5
A Cyclothymia	5, 4 with 3, 2, 1	t	·05	low
B General Intelligence ..	5 with 4, 3, 2, 1	C.R.	·05	high
E Dominance	5 with 4, 3, 2, 1	C.R.	·01	high
L Protension	5 with 4, 3, 2, 1	C.R.	·01	low
M Autia	5 with 4, 3, 2, 1	t	·05	high
Q1 Radicalism	5 with 4, 3, 2, 1	C R.	·01	low
Q3 High Self-Sentiment ..	5 with 4, 3, 2, 1	C.R.	·01	low

It would seem that the better teachers were not warmer and more sociable than the average teacher in that school, and that it was the average and below average teacher who tended to be more sociable. The superior teacher appeared to score significantly higher than all other teachers on the tests of general intelligence, dominance, relaxed security, introversion and individuality, and conservatism, but also obtained significantly lower scores in Q3 which has loadings from tests of socially approved character responses, self-control, persistence, foresight, consideration of others and conscientiousness. The Q3 – tendency appears surprising on first inspection; however, later in the Test handbook another definition was given in that high scores usually are taken to indicate " the extent to which the person has crystallised for himself a clear, consistent, admired pattern of socially approved behaviour to which he strives to conform "[2] (p. 19). It would seem that the superior teacher in this school did not strive to conform to a socially accepted pattern of behaviour.

M+ (Autia) individuals are said to " express significantly more dissatisfaction with the group unity "[2] (p. 16). This tendency could be checked on two related questions which appeared in the professional questionnaire. Only five teachers listed five or more causes of dissatisfaction and all these teachers scored 9 or 10 on Factor M. The question on degree of satisfaction with teaching as a profession showed no such discrimination and thus it would seem that emphasis might be on *expressing* rather than *having* greater dissatisfaction with the group.

Warburton, Butcher and Forrest[8] noted that the final teaching mark of graduates of a British University in their year of professional teacher training correlated with factors G (·236), I (·220) and Q3 (·223) and that the Personality Factors provided the best prediction of practical teaching mark, with a multiple correlation of ·438.

Lamke[5] with classroom observations combined with head teachers' opinions on thirty-two American teachers, found that good teachers were distinguished by F+, H+, B –, A+ scores, whilst poor teachers tended to have F –, H – scores. Lamke suggested that a pattern of response rather than individual responses were indicators of the quality of teachers. On the other hand, Schwartz,[6] using assessment of the teacher in training and in his second year in the profession, did not find a significant correlation between 16 P.F. scores and assesssed teaching mark.

When comparison by sex was made the mean profiles for the women and the men teachers appeared as in Fig. 3.

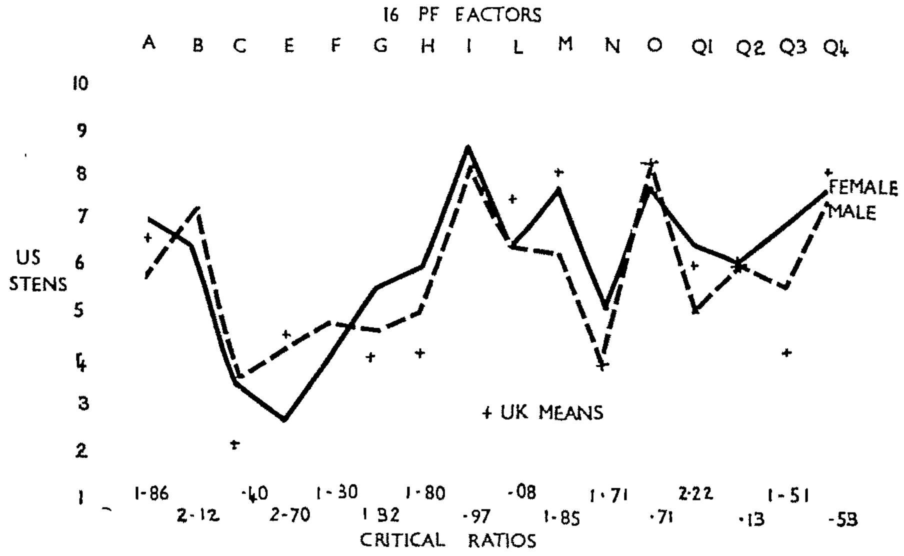


FIGURE 3

Sex differentiating factors (p < ·05, C.R.) were those of General Intelligence (B), Dominance (E) and Radicalism (Q1), and interesting, near-significant differences appeared in Factors A, H and M. As sex and rated teaching ability appeared to be linked as in Table 2, first order partial correlations were calculated between teaching ability and those 16 P.F. factors which appeared to offer discrimination between sex and/or teaching ability at the first analyses.

TABLE 2

DISTRIBUTION OF RATED TEACHING ABILITY AMONG THE MALE AND FEMALE TEACHERS.

	Teaching Ability Groups					Total	Biserial r
	5	4	3	2	1		
Male	7	3	5	2	2	19	·47
Female	0	3	6	5	2	16	

Partial Correlations.

Teaching ability was correlated with the relevant 16 P.F. factors by means of Pearson's ' r ' whilst biserial ' r ' was used for the dichotomous sex correlation with the selected 16 P.F. scores. The results are given in Table 3.

TABLE 3

CORRELATIONS AND PARTIAL CORRELATION OF SEX (2) AND TEACHING ABILITY (1) WITH SELECTED 16 P.F. FACTORS.

Variable	Statistic	16 P.F. Factors (x)							
		A	B	E	H	L	M	Q1	Q3
Teaching Ability (1) ..	Pearson r ..	−·21	·40*	·13	·10	−·39*	·09	−·26	−·21
Sex (Male +ve)	Biserial r..	−·38	·43	·52	−·26	·02	−·37	−·43	−·31
Teaching Ability	partial $r_{1x\cdot2}$	−·04	·25	−·15	·26	−·45*	·33*	−·07	−·08

* p< ·05.

It would, therefore, appear that, in the eyes of one head teacher, the best teachers differ from their less efficient staff colleagues in that they are relaxed, secure, accepting and adaptable, and less concerned with striving to conform to a socially approved pattern of behaviour. Such self-confidence and security might be the hallmark of success but the circular question arises as to whether these attitudes produce good teaching or are the result of experiencing success in the teaching situation.

One factor, Parmia *vs.* Threctia (H), failed to reach the figure for significance at ·05 level (C.R.=1·58). In the light of its theoretical likelihood, " distinguishing suitability for those occupations demanding ability to face wear and tear in dealing with people and gruelling emotional situations "[2] (p. 14) and its significance reported elsewhere,[3] perhaps it is worthy of further investigation in larger samples. The use of partial correlation to allow for sex differences in response might obscure a real difference in teaching ability between the sexes in the staff of this school. The rating distribution which placed seven men but no women in the highest teaching ability group might suggest either a real difference or that separate scales were, in fact, being used to assess the men and women in the staffroom. The significant difference between the mean scores of men and women in general intelligence would suggest that there is a real difference between this school staffroom and the general population. As regards the factor of resilience (H) and to a lesser extent general intelligence (B) (C.R.=1·53), partialling out the sex variable might have obscured a significant correlation with teaching ability.

Second Order Factors.

The second order factors of anxiety and extraversion were calculated for each subject and the mean scores for each of the five assessed teaching ability groups were calculated (Table 4).

TABLE 4

MEAN SECOND ORDER FACTOR SCORES FOR EACH OF THE FIVE ASSESSED TEACHING ABILITY GROUP.

	Teaching Ability Groups				
	5	4	3	2	1
Anxiety	7·51	7·54	7·64	7·99	6·72
Extraversion	4·99	3·26	4·28	4·35	3·99

Overall, the anxiety scores would appear high and the extraversion scores low. This could reflect differences between this group of teachers and the local population generally, but it is more probable that they are examples of U.K./U.S.A. differences on this American normed test.

Sub-group numbers were both low and unequal and the variation in individual scores within certain groups was quite marked. For example, the anxiety mean of 6·72 for Group 1 came from four scores 9·62, 9·48, 3·90 and 3·87. Using critical ratio, not one of the differences between anxiety means reached the 5 per cent. level of significance and only that between Groups 5 and 4 on the extraversion factor.

Rater Influence.

In utilising one person's rating of another it always seems advisable to know the rater's background, as work by Taft[7] suggests that individuals tend to make comparisons with themselves in making rating judgments. Some objective check of the rating in this study is possible as one of the other traits of the staff the Head Teacher was asked to evaluate was general intellectual ability. This is also tested by Factor B in the 16 P.F. Test. The intercorrelation between these two estimates was ·151.

The Head Teacher co-operated in personally completing the 16 P.F. Test to enable some comparisons to be made between these specific scores and the teaching efficiency ratings of the other members of the school staff. To see if any regular tendency existed, the personal score was compared with the mean scores of the two extreme groups rated on teacher efficiency. In nine cases the personal scores approximated closer to the top teaching ability group (+ve) whilst in seven to the lowest rated group (–ve). The nine +ve groups were A, B, E, F, H, L, M, N, Q3, and hence the best teachers are closer than their worst colleagues to the Head Teacher's scores on self-sentiment, shrewdness, unconventionality, protension, adventurousness, surgency, general intelligence and warmth of feelings. On eight factors—B, E, F, G, H, L, Q2, Q4—the Head Teacher's personal scores fell outside the range determined by the means of the five groups of teachers rated on teaching ability. Marginal differences existed on intelligence (B+), Dominance (E+, Surgency (F –), Character Strength (G –), Protension (L –) and High Ergic Tension (Q4 –), but differences in excess of one sten score occurred in Adventurousness (H – 2 sten) and Self-sufficiency (Q2+ 1·5 sten). Whether these are indicative of differences between Head Teachers and their assistants is something that might be worth investigating.

IV.—COMMENTS AND CONCLUSIONS.

In the literature criteria and samples have been dichotomous. Most often students under training have formed the experimental sample and final teaching mark the criterion score. Alternatively, practising teachers are assessed by classroom observation and by head teachers' opinions. There seems to be considerable lack of agreement between the findings with training and practising groups in teaching[9] and in general medical practice.[5]

The differences, particularly on Q3, between the findings of this study and those of Warburton, Butcher and Forrest may be the result of differences between student and practising teachers, between the criteria of head teachers and university education lecturers, or of random differences. Discrepancies between this and Lamke's work may be the result of cultural variations which required a more socially orientated teacher in the U.S.A. and a more academically inclined individual to fit the British education system. That might explain the polarity of the A+, F+ and perhaps B – in the American study and the A –, E+ and B+ in this study. Likely common ground might be on factor H, resilience, as whatever the cultural mores the contact, between the minds of forty children and that of one adult calls for some form of psychological shock absorber.

It would seem that important differences might exist between the ratings of student teachers by their training staff and those of practising teachers by their headmasters. Similarly, cross-cultural differences on the personality criteria of good teachers might exist between the U.S.A. and U.K. education systems.

This study would suggest that further work into the relation between the teaching ability of school staff as rated by the head teacher and 16 P.F. Test factors, A, H, L, and M could prove interesting.

V.—REFERENCES.

[1] BARR, A. S., *et al.* (1961). Wisconsin studies of the measurement and prediction of teacher effectiveness. *J. Exp. Educ.*, 30, 1-156.

[2] CATTELL, R. B, SAUNDERS, D. R., and STICE, G. (1957). *Handbook for the Sixteen Personality Factor Questionnaire.* Institute for Personality and Ability Testing, Champaign, Illinois.

[3] LAMKE, T. A (1951). Personality and teaching success. *J. Exp. Educ.*, 14, 75-100.

[4] OLIVER, R. A. C., and BUTCHER, H. J. (1962). Teachers' attitudes to education. *Brit. J. Soc. Clin., Psychol.*, 1, 56-59.

[5] PETERSON, O. L., ANDREWS, L. P., SPAIN, O., and GREENBERG, B. G. An analytical study of North Carolina General Practice, 1953-4. *J. Med. Education*, 31, 12 (ii), Dec., 1956.

[6] SCHWARTZ, A. N. (1950). A study of the discriminating efficiency of certain tests of the primary source personality traits of teachers. *J. Exp. Educ.*, 19.1, 63-94.

[7] TAFT, R. (1963). The judge as a variable in interpersonal judgments. In DAVID, H. P., and BRENGELMANN, J. C., Perceptions in Personality Research. Springer, New York, 1960.

[8] WARBURTON, F. W., BUTCHER, H. J., and FORREST, G. M. (1963). Predicting student performance in a university department of education. *Brit. J. Educ. Psychol.*, 33, 68-82.

[9] WISEMAN, S, and START, K. B (1965). A follow-up of teachers five years after completing their training. *Brit. J. Educ. Psychol.*, 35, 342-361.

(*Manuscript received 22nd June*, 1965)

THE RELATIONSHIP BETWEEN LEVEL AND SPEED ON TWO INTELLIGENCE TESTS

By AUBREY J. YATES

(*University of New England**)

SUMMARY. A group of first year engineering students was given the Progressive Matrices (1947) with initial and final time limits; the Nufferno Level test, with a time limit; and an objective (mathematical) test with initial and final time limits. As in a previous study, the Matrices test revealed a slow-working group which was handicapped by the time limit. On the Level test, this group completed fewer items within the time limit, but obtained as a high a Level score as the faster-working students. The group was shown to be handicapped by the time limit placed on the objective test. The implications of these results for university selection procedures, and the relevance of the personality dimension of introversion-extraversion as a possible explanatory factor for the group differences, are discussed.

I.—INTRODUCTION.

IN an earlier study (Yates, 1963) it was shown that the score obtained by some university students on Progressive Matrices (1947) would represent a significant underestimate of their intellectual level if the test were given with a time limit of 40 minutes. Analysis of the performance of these students during the time limit revealed a distinctive pattern—fewer items being attempted, but the majority of these being correctly solved. In other words, these students appear to work slowly but accurately. Indeed, even though handicapped by a time limit, the actual item pass score after 40 minutes was not markedly different from that of groups showing little or no increase in score when additional time is allowed. These ' slow ' students are, in fact, probably significantly brighter than the average university student. Although an analysis of examination results in Psychology I did not reveal any relationship between score increase on the Matrices and pass/fail examination results, it was pointed out that slow-working but bright students might well reveal enough knowledge in a verbal type examination to obtain a satisfactory mark; and that, of course, they might have obtained even better results had they been given more time. In more ' objective ' type examinations these students might be more severely handicapped.

The present study was designed to explore further three aspects of the phenomenon reported in the previous paper. Firstly, it seemed desirable to see whether a similar group of slow but highly intelligence students could be identified when the Matrices study was repeated on a new sample. Secondly, an attempt was made to verify that the handicap was one of speed rather than power by utilizing a different intelligence test which controlled for differences in speed of working. Thirdly, the performance of the ' handicapped ' group on a timed objective test was examined to see whether it was poorer than would be expected from the intellectual level of the subject.

---

* This study was carried out at the University of Western Australia. Grateful acknowledgement is due to Dr. H. Briner of the Department of Mathematics at that University, who constructed, administered and scored the objective test; and Dr. G. Stanley, who assisted with test administration.

II.—PROCEDURE.

A complete class of ninety-five first-year engineering students was tested. This class was homogeneous in that all students were taking the same set of courses and would continue to follow the same study programme throughout their engineering courses. Because the class was an engineering group, it contained a fairly wide range of mathematical competence, which was important from the point of view of the objective test. Seven subjects were rejected from the anlysis, six because of failure to follow instructions correctly in the Matrices and one because he had to leave before the completion of testing.

In the first testing session, the Matrices was presented with a thirty-minute time limit, at the end of which a circle was drawn round the most recently completed answer. A further fifteen minutes was then allowed. The group was informed of the thirty minute time limit before beginning the test, but was not informed of the additional fifteen minutes until the first time limit was up.

After completion of the Matrices, the Nufferno Level Test GL/2C 46 was administered (Furneaux, 1956), the instructions for administration being closely followed, and with a time limit of 25 minutes. The group was informed of the time limit and also that there would be no extension of time. Thus, both tests were administered under the same expectations as to time limit.

Several weeks later, during a scheduled mathematics period, the group was given an objective mathematics test by the lecturer in charge. The test consisted of short mathematical problems derived from the lecture series being given at the time and was administered with a total time limit of 45 minutes. However, as with the Matrices, the group was initially informed that the time-limit would be thirty minutes. At the end of the time they were asked to indicate the last question answered and then proceeded for a further fifteen minutes.

III.—HYPOTHESES.

The following hypotheses were formulated :

(i) Administration of the Matrices with initial and final time limits of thirty and forty-five minutes, respectively, would result in the identification (by analysis of score increase in relation to pass, fail, and not attempted item numbers in the first thirty minutes) of a group of students similar to that found in the earlier study.

(ii) The Nufferno Level test is constructed in cylcic form—that is, in each successive cycle of items, every difficulty level is represented, so that it does not matter how many cycles are completed within a given time limit. The only effect of completing more cycles is to increase the reliability of the level score. Hence, two predictions could be made in relation to this test: (*a*) the high-increase group on the matrices would complete significantly fewer items on the Level test within the given time limit, as compared to the low-increase group ; (*b*) the high-increase group on the Matrices would obtain the same Level score as the low-increase group within the given time-limit.
In other words, the Nufferno Level test should show clearly the slow rate but high level of the group identified on the Matrices test.

(iii) On the Objective mathematics test, the high increase group should obtain significantly lower scores within the first time limit, but this difference should be much less at the end of the second time limit.

IV.—RESULTS.

Hypothesis 1. In the previous study (Yates, 1963), one group started with item one, the other with item 21. Hence, they were compared on items 21–48 only. Similarly, in the present analysis, only these items were used. Subjects were grouped according to score increase in the extra fifteen minutes and compared with respect to average number of items passed, failed, and not attempted in the first thirty minutes. The same trends as in the previous study are evident although the differentiation of the groups is not quite as clear-cut, probably because, for unavoidable reasons, less time was spent on both the first part of the test (thirty as against forty minutes) and on the second (fifteen as against twenty minutes). Table 1 shows clearly that the students showing the greatest score increase in the additional time tend to work slowly but accurately, as indicated by the increase in number of items not attempted, and fewer items failed, in the first thirty minutes.

TABLE 1

AVERAGE NUMBER OF ITEMS PASSED, FAILED, AND NOT ATTEMPTED ON PROGRESSIVE MATRICES (1947) BY GROUPS CLASSIFIED BY SCORE-INCREASE IN ADDITIONAL TIME.

Score Increase	No. of Items			N
	Passed	Failed	Not Attempted	
0	13·9	10·6	3·6	24
1	13·9	7·8	6·3	19
2	13·6	5·8	8·6	21
3	15·4	4·3	8·4	11
4	12·8	4·7	10·5	6
5	13·8	3·4	10·8	5
6	12·0	2·0	14·0	2

Hypothesis 2*a*. Table 2 shows the relationship between number of Nufferno Level test items *not attempted* within the time limit and additional-time Matrices score increases. A clear trend is evident in support of the hypothesis that slow workers on the Matrices tend to be slow workers on the Level test as well. Score-increase groups (0, 1, 2, 3) obtained a mean Level not-attempted score of 8·71 (σ5·51) compared with a mean score of 12·15 (σ4·33) for the score-increase groups (4, 5, 6), the difference being significant ($t=2{\cdot}116$; $p<{\cdot}05$).

Hypothesis 2*b*. Table 2 also shows the average Level score for each score-increase group. The hypothesis that the slow workers would not be inferior on level score, when the time factor is controlled, is supported. For score-increase groups (0, 1, 2, 3) the mean Level score is 213·24 (σ56·80), while for score-increase groups (4, 5, 6) it is 210·15 (σ56·18), this difference being insignificant ($t=0{\cdot}174$).

It may be noted in passing that the standardization data presented in the Nufferno manual (based on 1,000 first-year Arts and Science students) places the fiftieth percentile at a score of 180.

Hypothesis 3. This hypothesis is only partially supported, but the results of analysis of the objective (mathematical) test are very interesting. There are no significant differences between the various groups in scores obtained on the first part of the test, but the Matrices score-increase groups are clearly not revealing the full extent of their ability during the first time limit to the same degree as the no-increase Matrices group. They tend to show larger score increases on the objective test in the additional available time in much the same way as they do on the Matrices, as shown by the final average scores.

TABLE 2

NUMBER OF ITEMS NOT ATTEMPTED AND LEVEL SCORE ON NUFFERNO LEVEL TEST IN RELATION TO MATRICES SCORE INCREASE.

Matrices Score Increase	Level Test Items Not Attempted	Level Score	N
0	6·90	205	24
1	8·21	219	19
2	11·38	191 } 224	21
3	8·47	263	11
4	13·66	211	6
5	9·80	214 } 200	5
6	13·50	198	2

TABLE 3

RELATIONSHIP BETWEEN OBJECTIVE TEST SCORES AND MATRICES SCORE INCREASE.

Matrices Score Increase	Objective Test Scores		N
	Initial	Final	
0	15·71	21·79	24
1	16·06	23·06	18*
2	17·38	24·52	21
3	16·27	22·45	11
4	18·83	28·00	6
5	16·40	24·00	5
6	14·00	23·50	2

* Scores for one subject were not obtained.

V.—DISCUSSION.

Although dealing with a fairly homogeneous sample of first-year engineering students above the general intellectual average of Australian university students, the results obtained in the present study both confirm and extend previous findings (Yates, 1961, 1963). Firstly, a test like Progressive Matrices (1947), which confounds level and speed of intellectual performance, if given with limited time available, is likely to markedly underestimate the intellectual capacity of a small, but significant, group of students. The validity of this finding is further demonstrated by the use of the Nufferno Level test in which, as expected, this small group completes fewer items in a given time limit, but, because slowness of working is not a handicap, obtains level scores the equal of faster workers.

Thirdly, whereas a previous study (Yates, 1963) was unable to demonstrate that this group was handicapped in written (non-objectively scored) examinations, the present study shows that the speed handicap does show up on a more objective (arithmetic) test, given with rather severe time limits.

This study, and the previous ones, do not, of course, throw any light on the reasons for slowness accompanied by high intellectual capacity, and merely suggest that great care be taken in interpreting results of intelligence tests given with time limitations. There are, however, important suggestions in the literature relating personality factors to 'style' of performance on intelligence tests. The most important dimension of personality which might be relevant to the results reported here is introversion-extraversion (Eysenck, 1947). The finding that extraverts tend to work quickly but inaccurately, whereas introverts tend to work slowly and accurately suggests that group (O) inthis investigation would obtain a higher extraversion score than groups (4, 5, 6). If this were so, then *both* groups would be failing to reveal their true intellectual capacity on the Matrices, but for different reasons. This suggestion should, however, be regarded very tentatively, since the interaction between personality factors and level and speed of intellectual functioning is very complex (Lynn and Gordon, 1961; Eysenck and White, 1964).

Finally, it should be noted that a good deal of selection has already taken place by the time of entry to university. The possibility cannot be ignored that many slow but highly intelligent students are eliminated from consideration for entry to a university because of the stress on both speed and power in selection procedures. This is not to suggest that speed of working is not a relevant factor in deciding university entry. However, it is possible that some of these students might well be able to profit from a university education under favourable conditions.

VI.—REFERENCES.

EYSENCK, H. J. (1947). *Dimensions of Personality.* London: Routledge and Kegan Paul.

EYSENCK, H. J., and WHITE, P. O. (1964). Personality and the measurement of intelligence. *Brit. J. Educ. Psychol.*, 34, 197-202.

FURNEAUX, W. D. (1956). The Nufferno Manual of Level tests. (Published by the Institute of Psychiatry and distributed by the National Foundation for Educational Research.)

LYNN, R., and GORDON, I. E (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 194-203.

RAVEN, J. C. (no date). *Progressive Matrices* (1947), *Sets I and II.* Published privately by the Crichton Royal.

YATES, A. J. (1961). Item analysis of Progressive Matrices (1947). *Brit. J. Educ. Psychol*, 1961, 31, 152-157.

YATES, A. J. (1963). A further study of Progressive Matrices (1947). *Brit. J. Educ. Psychol.*, 1963, 33, 307-311.

(*Manuscript received 28th June,* 1965)

THE EFFICACY OF THE COLOURED PROGRESSIVE MATRICES AS A GROUP TEST WITH YOUNG CHILDREN

By P. S. FREYBURG
(*Massey University of Manawatu, N.Z.*)

SUMMARY. The Coloured Progressive Matrices test has proved to be a useful non-verbal instrument apart from some doubts concerning its reliability. A pilot investigation gave reason for belief that the test could be group-administered with higher reliability than had hitherto been reported. The test was administered to 159 6 to 7-year-olds, coefficients of internal consistency and stability being ·89 and ·87—·76, respectively. Relationship of matrices scores with those obtained from other instruments are quoted, together with score distributions and directions for group administration with young children.

I.—INTRODUCTION.

TEACHERS and research workers have frequently expressed their need for an efficient group test of general intellectual ability suitable for administration with the 6–8-year age group. Unfortunately, the reliability co-efficients of existing group tests designed for use with such children have, in general, proved to be rather lower than is desirable if scores of individuals are to be compared. This handicap has been particularly evident with regard to non-verbal tests, whether administered as group instruments or individually; yet a good non-verbal (or, at the least, a non-reading) test of intelligence would be invaluable at this age level for diagnostic purposes in cases of scholastic retardation.

While Raven (1948) does not consider his Progressive Matrices test by itself an instrument for measuring intellectual ability, its saturation with ' g ' has been reported to be as high as ·8 by Slater (1948) and ·82 by Burt (1954). The two main forms of the test (Sets A–E and A, Ab, B) have been used extensively for clinical purposes, as Burke's review article (1958) shows. Burke has expressed himself satisfied with " the abundant evidence of its concurrent validity " as a test of what is commonly known as intelligence.

The Coloured Progressive Matrices (Sets A, Ab, B), which Raven devised for use with young children, would at first glance, thus appear to be a promising instrument for administration in the early school years, except for two considerations: the comparatively small and geographically limited sample on which the published norms are based, and the low test-retest correlations of ·6 for 6-year-olds and ·8 for 9-year-olds quoted by Raven (1960), even when the test was given individually. The first of these objections is the less important one, as the test could still be useful for ranking children within a class or school, and in any case a more thorough standardization could be undertaken without much difficulty. If the reliability of the test is really as low for young children as Raven has stated, however, its use would rarely be justified.

There are some grounds for believing, nonetheless, that Raven's quoted reliability co-efficients may be underestimates: certainly the figures were obtained from testing rather small groups (n=58 and 61, respectively). For instance, Martin and Wiechers (1954) have reported a correlation of ·91 between Coloured Progressive Matrices (CPM) scores and WISC I.Qs. for 100 9-year-old children, which would be unlikely if either of the tests had a high proportion of error variance. And Tuddenham, *et al.* (1958) obtained internal consistency co-efficients (Kuder-Richardson) ranging from ·71 to ·90, with a median of ·83 for a total of 168 6 and 7-year-olds.

The matrix type of item has recently been found useful in the investigation of concept formation and classification (Inhelder and Piaget, 1964; Bruner, 1964), apart from any value it may have in tests of general intellectual ability. The inclusion of the CPM in the longitudinal study of cognitive development in young children to be undertaken by the writer, thus seemed an obvious choice but for the questions of reliability and suitability for group administration.

Sperrazo and Wilkins (1958) had already found that the CPM could be successfully administered as a group test by means of coloured slides with 9 to 11-year-olds. Tuddenham, *et al.* (1958) used a mimeographed black and white version of the CPM in their study. Internal consistency and test stability are by no means equivalent, however, although measurements of both are commonly referred to as indices of reliability; and it is the self-consistency of the test which has received most attention up till now. Nevertheless, a pilot experiment revealed that the C.P.M. could, indeed, be used as a group test with a reasonable degree both of self-consistency and stability,and a more rigorous test of its usefulness was, therefore, undertaken.

II.—TEST MODIFICATIONS AND SAMPLE POPULATION.

The main problem in adapting the test for group administration was to get 6 to 8-year-old children to record their answers in the correct place on a separate answer sheet. By the middle of the second year in the infant school most children can at least read numbers up to 12 and write the numbers 1–6, which was all that would be required in this respect. The published answer sheet for use by older children has a simple three-column layout, so that children can write their answers directly alongside the numbers of the twelve items in each set. It was found that even the dullest 6-year-olds tested had little difficulty in recording their responses correctly after guidance with the first three items.

Raven's (1960) instructions for administering the book form of the test were modified as little as possible. The children were given ample time to answer each item, but were instructed not to turn to the next page until told to do so. The possibility of misplacement of answers was minimized in other ways, also—by pointing out the colour of each new pattern to ensure that children did not turn over two pages at once, and by reminding them explicitly to put their answers alongside the number in the appropriate column. By moving around the room fairly frequently, the tester could readily prevent misplacements but it was necessary to tell children to put a dash (' line ') if they could not decide which was the right answer. The modified instructions are given in detail in an Appendix below. It is suggested that no more than twenty children should be tested at one time.

The sample population for which results will be given consisted initially of 159 children aged 5 : 9 to 7 : 10 with a median age of approximately 6 : 9 years. These were drawn from a total of nine classes in four schools in a relatively prosperous New Zealand inland urban centre with a total population of about 42,000 people. Although socio-economic differences between the areas in which each school is situated are not marked, such differences as do exist were taken into account in order to avoid bias in the sample. Children who were, after more than a year's schooling, complete non-readers were excluded from the sample, which may thus represent a distribution slightly curtailed at the lower levels of intelligence. The range of intelligence quotients (Thurstone's Primary Mental Abilites Test, 5–7 year level) was from 77 to 133, with a mean of approximately 104, using the United States norms.

III.—RESULTS.

(a) *Internal Consistency.*

For 159 children the corrected split-half (odd-even items) correlation co-efficient was ·896. The same data produced a reliability coefficient of ·894 when analyzed to take into account differences in variance between the two arbitrary halves into which the test scores were divided. The test is thus a fairly homogeneous one, even for young children.

(b) *Stability.*

The pilot investigation made use of three classes with median ages of 6 : 6, 7 : 4 and 8 : 4 years. The test-retest correlations, with an interval of a week between administrations, were respectively ·87, ·83 and ·81. For the total number of ninety-eight children in these three classes the co-efficient was ·871.

From the initial group of 159 children, seventy-five took the CPM and other tests again three months after the first administration. The test-retest correlations were ·766 for the CPM, ·765 for the Primary Mental Abilities 5–7 form, and ·817 for a seventy-five-item objective test of concept development (along Piagetian lines) which had been constructed expecially for the investigation. An experimental group of forty-seven children completed the test at three-monthly intervals over a period of two years, the CPM correlation between initial and final testing being ·599. A control group of fifty-four children which took only the initial and final tests had a retest correlation of ·466. In both the latter cases there was some evidence of a ceiling effect which would have tended to lower the correlations.

(c) *Distribution of Scores.*

The pattern of CPM scores obtained by 151 children who completed all tests in the battery administered at the initial testing is given in Table 1.

Bearing in mind the restricted nature of the sample, these figures should not be regarded as norms but are presented simply to show that group administration of the test is feasible. The raw scores obtained by 6-year-olds were higher than those quoted at various percentile levels by Raven (1960) using an individual administration, but correspond fairly well with the results obtained by Green and Ewert (1955) in the United States. It should be noted that the tests were given about two-thirds of the way through the children's second infant-school year.

Distributions are given separately for boys and girls, as a significant sex difference ($t=3{\cdot}24$; $p<{\cdot}01$) was found between the means for the two sexes. This is contrary to the findings of some United States investigators (Tuddenham,

TABLE 1

RAW SCORES OBTAINED BY SIX-YEAR-OLDS AND SECOND YEAR INFANTS EXPRESSED AS PERCENTILE RANKS.

Raw Score	Age group 6 : 0—6 : 11			Second year infant group (age range 5 : 9—7 : 10 years)		
	Boys n=45	Girls n=63	Together n=108	Boys n=64	Girls n=87	Together n=151
33+	99		99	97		99
32						
31	97		98	96		98
30	95	99	97	94	99	97
29	92	96	94	91	97	95
28	90	94	93	90	96	93
27	88		89	87		90
26	84	90	87	80	92	87
25	79	88	84	73	90	83
24	77 a	86	82	68	89	80
23	68	80	....75....	60	84	74
22	62	72	68	56	79	69
21	53 b	68	61	48	73	63
20	46	66	58	42	67	57
19		60	52	39	59	51
18	42	50	47	37	51	46
17	37	46	42	31	48	41
16	31 a	39	36	26	41	34
15	24	33	29	20	35	29
14	16	....25 ...	21	14	27	22
13	9	11	10	6	13	10
12	4	5	5	3	7	5
11						
10	2	3	3	2	5	3

a—Quartiles. b—Median.

et al., 1958 ; Higgins and Sivers, 1958). There was no evidence of any significant differences according to socio-economic background of the children, concerning which there have been conflicting results in American research using the CPM.

(*d*) *Relationship with Other Tests.*

The correlations between CPM scores and those obtained by children on other tests administered at the same time, are set out in Table 2. Only the scores for the control group have been used at the 8-year-old level, because of the possible influence of differential practice effects in the case of the other children.

TABLE 2

CORRELATION OF C.P.M. SCORES WITH OTHER COGNITIVE TEST SCORES AT TWO AGE LEVELS.

Test		Median age 6 : 10 n=151 r (SE_r)	Test		Median age 8 : 10 n=54 r (SE_r)
PMA 5–7 Years Factor Scores	V	·176 (·08)	PMA 7–11 Years Factor Scores	V	·242 (·13)
	P	·277 (·08)		P	·345 (·12)
	S	·486 (·06)		S	·470 (·10)
	Q	·377 (·07)		R	·439 (·11)
				N	·308 (·12)
PMA 5–7 Weighted MA		·463 (·06)	PMA 7–11 Weighted MA		·373 (·12)
Concept Development		·407 (·07)	Concept Development		·550 (·09)

These low to moderate correlations are in agreement with those reported earlier for both the 1938 and 1947 forms of the Progressive Matrices, which have been surveyed in detail by Burke (1958). The highest correlation for 6-year-olds was between the CPM and PMA Space Factor Scores; the difference between this relationship and that between the CPM and Perceptual-Speed (P) Factor in Thurstone's test, which is statistically significant for correlated samples ($t=3{\cdot}55$, $p<{\cdot}01$), indicates that the matrices test involves more than just perceptual skills at this age level. The relationship between the 1938 form and tests of spatial abilities has already been noted by Emmett (1949).

IV.—DISCUSSION AND CONCLUSIONS.

It is somewhat unusual, to say the least, that later research work should bolster a claim for a higher degree of reliability for a test than has originally been put forward by its author, especially when the method of administration is changed from an individual to a group basis. It may well be, however, that the group administration of a comparatively straightforward instrument provides a more standardized situation than does individual testing. Even the most conscientious tester can unconsciously introduce variations in his procedures where he is testing a succession of individuals!

One confounding factor in assessing the reliability of cognitive tests with this age group may well be intra-individual variability. Several writers—e.g., Garrett (1946), Hsu (1948), Burt (1954) and Vernon (1961)—have suggested that cognitive abilities appear on the evidence to be relatively undifferentiated before 9–12 years of age, and Cronbach (1960) attributes some of the error variance in intellectual measurements at earlier ages to resulting instability in behaviour patterns. Thus, when stability of measurement is estimated (test-retest or equivalent form reliability) it turns out to be disappointingly low compared with measures of internal consistency derived by the use of split-half or the Kuder-Richardson methods.

Nevertheless, some individual tests commonly used with 6 and 7-year-olds can claim a fairly high degree of stability. Terman and Merrill (1937), for instance, obtained a correlation of ·91 for an unselected group of 7-year-olds tested with the Revised Stanford-Binet Forms L and M administered within a few days of each other. A test with heterogeneous items would, of course, be likely to yield more stable measurements because the final score represents an average of the results of sampling several kinds of ability. At the same time, a heterogeneous test will require more adjustment to differing content than would a homogeneous type of instrument; hence, the need for the flexibility of an individual administration.

The Progressive Matrices test has proved to be relatively homogenous in content, and the fact that it can be successfully administered to small groups of young children is, therefore, perhaps not so surprising after all. Furthermore, it seems probable that improvements in both internal consistency and stability could be achieved without much difficulty by re-designing some of the items, following more stringent item analyses than have apparently been undertaken by Raven. Analyses by Banks and Sinha (1951) and Jordan and Bennett (1957) have already revealed deficiencies in the structure of sets A, Ab and B, in particular.

Although the Progressive Matrices can successfully differentiate between children of high, average and low general ability, it is still not clear exactly what the test is measuring. Estimates of the test variance which can be accounted for by ' g ' vary considerably (Banks and Sinha, 1951). At the same time, the results of the present investigation, amongst others, show that the CPM form has the potential of being a useful group test, which could provide valuable information regarding the cognitive development of young children. It should, therefore, repay further refinement, especially as a research instrument.

ACKNOWLEDGEMENTS.—The author acknowledges with gratitude the co-operation of the staffs of the schools concerned with this investigation, and the financial support of the N.Z. University Research Committee. The advice of Dr. A. E. Fieldhouse (Victoria University of Wellington) is also appreciated.

V.—REFERENCES.

BANKS, C., and SINHA, U. (1951). An item-analysis of the Progressive Matrices Test and Binet. *Brit. J. Psychol., Statist. Sect.*, 4, 91-94.

BRUNER, J. S. (1964). The course of cognitive growth *Amer Psychologist*, 19, 1-15

BURKE, H. R. (1958). Raven's Progressive Matrices: a review and critical evaluation. *J. Genet. Psychol.*, 93, 199-228

BURT, C. (1954). The differentiation of intellectual ability. *Brit. J. Educ. Psychol*, 24, 76-90.

CRONBACH, L. J. (1960). *Essentials of Psychological Testing.* New York: Harper.

EMMETT, W. G (1949). Evidence of a space factor at 11+ and earlier. *Brit. J Psychol., Statist Sect*, 2, 3-16.

GARRETT, H. E. (1946). A developmental theory of intelligence. *Amer. Psychologist*, 1, 372-378.

GREEN, M. W., and EWERT, J. (1955). Normative data on Progressive Matrices (1947). *J. Consult. Psychol.*, 19, 139-142.

HIGGINS, C., and SIVERS, C. H. (1958) A comparison of Stanford-Binet and Coloured Raven Progressive Matrices I Qs for children with low socio-economic status. *J Consult Psychol.*, 22, 465-468.

HSU, E. H. (1948). Factor analysis, differential bio-process and mental organization. *J. Gen. Psychol.*, 38, 147-157.

INHELDER, B., and PIAGET, J. (1964). *The Early Growth of Logic in the Child.* London: Routledge.

JORDAN, T. E., and BENNETT, C. M (1957). An item analysis of the Coloured Progressive Matrices (1947). *J. Consult. Psychol.*, 21, 222.

MARTIN, A. W., and WIECHERS, J. E. (1954). Raven's Coloured Progressive Matrices and the Wechsler Intelligence Scale for Children. *J. Consult. Psychol.*, 18, 143-144.

RAVEN, J. C. (1948). The comparative assessment of intellectual ability. *Brit. J. Psychol.*, 39, 12-19.

RAVEN, J. C. (1960). *Guide to Using the Coloured Progressive Matrices.* London: Lewis.

SLATER, P. (1947). The associtaion between age and score in the Progressive Matrices Test. *Brit. J. Psychol., Statist. Sect.*, 1, 64-69.

SLATER, P. (1948). Comment on 'The comparative assessment of intellectual ability.' *Brit. J. Psychol.*, 39, 20-21.

SPERRAZO, G., and WILKINS, W. L (1958). Further normative data on the Progressive Matrices. *J. Consult. Psychol.*, 22, 35-37.

TUDDENHAM, R. D., DAVIS, L., DAVISON, L., and SCHINDLER, R. (1958). An experimental group version for school children of the Progressive Matrices. *J. Consult. Psychol.*, 22, 30. (Also Tuddenham, R. D., personal communication.)

VERNON, P. E. (1961). *The Structure of Human Abilities.* London: Methuen (2nd. ed.).

(*Manuscript received* 21*st June*, 1965.)

VI.—APPENDIX.

Instructions for administration of the Coloured Progressive Matrices as a group test.

1. Give out answer forms and booklets.

2. Say: " First, write your name on the paper I have given you. Do not make any marks in the book."

3. Say: " Now, open the book at the first page. Look at the top part. It is a pattern with a piece left out of it. Each of the pieces at the bottom of the page is the right shape to fit in the top part, but only one of them would make the right pattern. Number 1 is the right shape, but it is not the right pattern. Number 2 is not a pattern at all. Number 3 is quite wrong. Number 6 is nearly right but the white piece shouldn't be there. Only one piece is right—which is it? "

Obtain the answer, and tell children to put the number of the answer beside Number 1 in the first (A) column. Check that all have done this.

4. Say: " Now turn to the next page, A 2. It is an orange pattern. Which piece would fit in and make the right pattern this time? Look at each piece in turn and then put the number of the right piece beside Number 2 on your paper, below your last answer."

Obtain the answer from one of the children and check as before that all children have recorded the answer in the correct place. Give further explanation where necessary.

5. Repeat this procedure with item A 3 but warn children not to tell anyone else what they have put, nor to look at anyone else's paper. When they have put down the answer—

Say: " Now, look at the piece you have chosen and make sure that it really would make the right pattern."

Obtain the answer and check as before, warning children that " we won't talk about any more of the answers for a while."

6. Proceed with the test, reminding children at every third problem to look again and see that they have chosen the right piece to make the pattern. As each page is turned, refer to the number of the problem and its colour.

7. As the test proceeds, check answer forms occasionally to make sure that they are being correctly filled in, taking particular care that the children start at the top of the row again for sets Ab and B. Obtain the answer for Ab 1 and B 1, and check as before, but no other assistance should be given with any other item.

8. Allow approximately 1 minute per item, although no time limits are imposed. Proceed to next item when all of the children have recorded an answer.

THE RELATIONSHIP BETWEEN PERSONALITY CHARACTERISTICS AND SCHOLASTIC SUCCESS IN ELEVEN-YEAR-OLD CHILDREN

BY JAMES RUSHTON
(*Department of Education, The University, Manchester*)

SUMMARY. Teacher ratings of personality and ability on 458 children aged 11 years were obtained as were their scores on a personality questionnaire and tests of cognitive ability. The scores were intercorrelated. The major finding was that, according to the C.P.Q. test and teacher ratings, well adjusted, extravert children have higher scholastic attainment.

I.—INTRODUCTION.

FOR a number of years the idea that personality factors play an important part in scholastic success has been accepted. There remains the problem of identifying and measuring the factors which lead to success or failure. Standardised questionnaires provide a useful instrument for assessing large samples and the Children's Personality Questionnaire (Cattell's C.P.Q. test) appears to be one of the more promising instruments available for use with children because :

(*a*) It provides measures on fourteen primary factors (the second order factors—anxiety and extraversion—are derived from weighted sums of the first order factors).

(*b*) The factors have been derived from a considerable body of research and they correspond broadly to those found in adults.

The majority of studies published to date are concerned with the two broad personality traits of anxiety-adjustment and extraversion-introversion. Nearly all are based on samples of adults or older children ; a good proportion are American and much of the evidence is conflicting. The evidence arraigns itself in this way. Of the twenty-seven researches so far reported which study the relationship between anxiety (neuroticism) and academic achievement, eleven researches favour the hypothesis that anxiety (neuroticism) is positively connected with academic achievement and sixteen suggest the contrary. Considering only studies based on the questionnaires of Cattell, Eysenck and Guilford, we find six out of fifteen linking neuroticism with achievement. The over-riding impression is that whenever the training course is particularly severe, neuroticism tends to be a positive influence. In other words, ' tensed up ' people are more able to deal with a tense situation.

However, if the evidence is kept strictly to work done with children, eleven out of thirteen studies connect stability, as opposed to anxiety, with ability, or seven out of eight if work not based on standardised questionnaire measures is excluded. Thus, in twenty-eight instances out of forty-two (or some 70 per cent. of the researches studied) stability or adjustment is positively connected with academic achievement.

Turning to the relationship between introversion-extraversion and scholastic performance, we find seventeen out of nineteen showing a significant connection between introversion and scholastic success. It must be stressed, however, that most of these researches are with older children or adults. Both Warburton (1962) and Butcher (1963) are of the opinion that research indicates that in junior school children extraversion may be a positive influence in scholastic success. Banks (1963) supports this theory.

The experiment was designed with the following aims in view :

(*a*) To investigate the relationship between personality assessments as measured on an objective questionnaire and cognitive ability measures as used in an actual 11 plus examination.

(*b*) To study the relationship between teacher ratings of personality and ability and the objective personality questionnaire.

II.—METHOD.

A sample of 458 boys and girls, aged 10–11 years, was drawn from fourteen county primary schools approximately representing the different socio-economic strata and school size in a large county borough. They also had, without exception, verbal reasoning quotients of 105 and above.

(*a*) The children were tested on each of the following tests :

- (i) Children's Personality Questionnaire (Cattell) Forms A and B.
- (ii) Moray House Verbal Reasoning Test No. 63.
- (iii) Moray House Arithmetic Test No. 30.
- (iv) Moray House English Test No. 30.
- (v) Moray House Spatial Test No. 2.
- (vi) A Teachers' Rating Scale of Fourteen Personality and Ability Traits.

(*b*) Treatment of Data.

The data used in the experiment were of four types :

- (i) Personality scores from the children's Personality Questionnaire. These data consisted of thirteen first order factors (factor B intelligence was withdrawn because the Moray House Verbal Reasoning Quotient is an adequate measure of this ability) and two second order factors.
- (ii) Ability measures on the Moray House battery.
- (iii) Teacher ratings on a Fourteen-Point Scale. For each of the fourteen traits a raw score total was converted to a standard score incorporating a correction for differences between age groups.
- (iv) Pass-Fail.

All of the above scores were converted to normalised ' t ' scores because many of the teacher ratings were heavily skewed in a negative direction. Means, standard deviations and intercorrelations were then calculated. All calculations with the exception of conversion to ' t ' scores were carried out on the Atlas Electronic Computer of the University of Manchester.

III.—RESULTS.

(*a*) *The relationship between cognitive ability measures and the personality variables from Cattell's C.P.Q.*

Table 1 below shows the product moment coefficients of correlation between the cognitive ability measures, the pass-fail dichotomy and Cattell's personality measures.

TABLE 1

CORRELATIONS BETWEEN CATTELL'S PERSONALITY VARIABLES, COGNITIVE ABILITIES AND PASS–FAIL DICHOTOMY.

	Verbal Reasoning	Arithmetic	English	Spatial	School Record	Pass *vs.* Fail
A+ Sociability						
C+ Ego Strength	145	144	149	124*	202	—219
D+ Excitability						
E+ Dominance		145			096*	
F+ Surgency	148	121*	160		112	—142
G+ Conscientiousness	180	164	135*		183	—132*
D+ Adventurousness		115*				
I+ Sensitivity				134*		
J+ Passive Individualism						
N+ Shrewdness						
O+ Guilt Proneness						
Q3+ Self Control					119*	
Q4+ Tenseness			—102*			
A—A Anxiety	—200	—149	—162	—103*	—195	147
E—I Extraversion	161	110*	130*		145	—138
Nr Neuroticism	—143	—153	—141		—157	167

*—Denotes significance at the 5 per cent. level.

N=458. (All the correlations shown are significant at the 5 per cent. or 1 per cent. level.)

The following points ensue :

(i) The second order factor of anxiety clearly has a negative correlation with verbal reasoning, arithmetic, English and school record and failure on the pass-fail dichotomy. This implies that the less anxious better adjusted child is most likely to succeed in school work at this age.

(ii) Extraversion, the other major second order factor, correlates positively with verbal reasoning quotient, school record, arithmetic and English and success (on the pass-fail dichotomy). Thus, at this age, extraversion favours scholastic success.

(iii) The ' experimental ' second order factor of Neuroticism follows the pattern of anxiety fairly closely. It is clear that low neuroticism makes for success in school just as does low anxiety.

(iv) The primary factors of Ego Strength, Surgency and Conscientiousness appear to assist all work in the cognitive field.

In summary, the stable extravert child is the one who, on these results, succeeds at his school work.

(*b*) *The relationship between Cattell's Personality Variables and Teacher Ratings of Personality and Ability.*

TABLE 2

CORRELATIONS BETWEEN CATTELL'S PERSONALITY FACTORS AND TEACHER RATINGS OF SCHOLASTIC AND PERSONALITY TRAITS. TOTAL AND SEX.

	Ambition 1	Alertness 2	Accuracy 3	Industriousness 4	Method 5	Speed of Working 6	Leadership 7	Relationship with Seniors 8	Relationship with Contemporaries 9	Co-operativeness 10	Self-consciousness 11	Perseverance 12	Reliability 13	Emotional Stability 14	Special Abilities, Arts, Crafts, etc. 15
A+ Sociability	096*			132*	115*	114*		098*	121*		097*				
C+ Ego Strength	174	165	181	135*	121*	189	165	245	121*	144	167	166	162		193
D+ Excitability		–124*		–116*					–120*				–130*		
E+ Dominance															
F+ Surgency	136					205	098*	098*					098*		
G+ Conscientiousness	190	200	191	208	178	103*	209	102*	103*		098*		170		
H+ Adventurousness	095*	104*		107*		162	115*					153			
I+ Sensitivity		094*	140	191	135*		136		102*		177		128*		
J+ Passive Individualism	–091*	–100*										103*			
N+ Shrewdness			–096*										–120*		
O+ Guilt Proneness		–101*													
Q3+ Self Control			136	112*			–100		114*		117*		118*		
Q4+ Tenseness							–100*								136
A–A Anxiety	–172	–195	–190	–193	–114*	–123*	–159	–131*	–129*		–146	–106*	–181		
E–I Extraversion	142	107*		114		135*		138			113*	169			
Nr Neuroticism	–151	–162						–137				–093*			

*Denotes Correlation is significant at the 5% level.
(Note the decimal points have been left out of this table.)

N=458. (All the correlations are significant at the 5 per cent. or 1 per cent. level.)

The outstanding features of this table are :

(i) On the second order factor of anxiety, children with low scores are rated by teachers as more fitted for an academic secondary education than others.

(ii) Extraversion correlates positively with teacher ratings of ambition, relations with seniors and perseverance. On these three important aspects the more extraverted children are rated by their teachers as being well fitted for an academic secondary education.

(iii) The second order factor of neuroticism confirms the finding that the more anxious children are less able academically.

(iv) The first order factors are confirmed to some extent by the teacher ratings. In particular, Ego Strength (C+) as measured by the questionnaire is a trait which teachers pick out, under different names of the fourteen traits, as an impoortant factor in scholastic success. In a similar way, conscientiousness (G+) self control (Q_3+) and sensitivity (I+) were also selected as traits which have a great influence on school success.

(v) It is also striking that children who rate themselves as surgent and enthusiastic, are rated by their teachers as ambitious and quick at their school tasks. Adventurousness also correlates positively with speed of working and perseverance. Tenseness has a negative connection with leadership, as is *prima facie* plausible. As in earlier researches, the coefficients of correlation in both of these tables are low as compared with correlations between ability measures.

III.—DISCUSSION.

To a teacher or educational psychologist, it seems obvious that at given levels of intelligence the well adjusted child should have a superior academic record to the anxious child. One of the findings of this research is quite definitely that in children of 11 years of age stability is positively correlated with academic success. This finding conflicts with those of Furneaux (1957) and Lynn (1959), but they were working with university students. It agrees, however with the findings of Hallworth (1961), Butcher, Ainsworth and Nesbitt (1962), Lunzer (1960) and Callard and Goodfellow (1962). Explanations of the positive association between stability and academic achievement could be that well adjusted children are not distracted to the same degree by their personal problems. Also that younger children are not affected to the same extent by personal sensitivity and ambition which could lead to anxiety, drive and success in older pupils and students.

Another finding was that extraverted children were more able scholastically than others. This finding is in direct conflict with most other researches in this field. Furneaux (1957), Lynn (1959), Cattell (1960) (*a*) and Holmes (1960) all found that introversion was of positive help in academic pursuits. Banks (1963) supports the writer's view. It is feasible that the thoughtful consideration of problems is too fragmentary to be of any value scholastically or that extraverted children make contacts more easily and become socially and mentally mature earlier, thus adapting themselves to the requirements of the classroom. With older children, however, one would expect the hard thinking, well read introvert to come into his own.

It would be interesting to see if the scholastically superior children who are extraverted at the age of 11 are still classed as extravert at say the age of 15 or whether bright extraverts at age 11 tend to become bright introverts at later ages.

It is noteworthy that teachers indirectly relate extraversion and stability to school attainment as their ratings correlate positively with the questionnaire factors. Thus, the academically successful child both rates himself and is rated by his teachers as extraverted and stable, although neither the child nor the teacher is consciously aware of the importance of these two major factors.

IV.—CONCLUSIONS.

(1) According to the C.P.Q. test the well-adjusted, extraverted children tend to have higher scholastic attainment as assessed by Moray House attainment tests.

(2) With regard to the correlations between the teacher ratings of personality, the main results confirm the first finding since it is the stable extravert children who have superior attainment.

In respect of stability the major findings in terms of primary factors are :

(i) that emotional maturity (C+) at this age would appear to assist children in scholastic pursuits.

(ii) That the more relaxed (Q4 –) children tend to be better at English.

(iii) That perseverance (G+) is positively related to success in school work at this age.

In respect of extraversion the major findings are :

(i) That teachers tend to rate children who are more ' easy going ' (A+) as more suitable for an academic career than others.

(ii) That the more dominant (E+) children tend to be more proficient at arithmetic.

(iii) That ' happy-go-lucky' (F+) children appear to be selected as more likely to benefit from an academic secondary education.

(3) The main relationship between the personality questionnaire results and the ratings of personality made by teachers show :

(i) That internally restrained (G+) children are rated by teachers as more persistent.

(ii) That children with better self-control (Q_3+) have better school records.

V.—REFERENCES.

BANKS, J. (1964). *The Relationship between Problem Solving in Arithmetic and Correct Attainment, Intelligence and Personality Characteristics in Junior School Children.* M.Ed. Thesis, University of Manchester.

BUTCHER, H. J., AINSWORTH, M. E., and NESBITT, J. E. (1963). Personality factors and school achievement—a comparison of British and American Children. *Brit. J. Educ. Psychol.*, 33, 276-286.

BUTCHER, H. J. (1963). A study of attainment, personality and motivation in school children. Unpublished report.

CALLARD, M. P. and GOODFELLOW, C. L. (1962). Three experiments using the Junior Mawdsley Personality Inventory. Neuroticism and extraversion in school boys as measured by J.M.P.I. *Brit. J. Educ. Psychol.*, 32, 241-251.

CATTELL, R. B. (1946). *Description and Measurement of Personality*. New York: World Book Company.

CATTELL, R. B., and PORTER, R. B. (1959). *Handbook for the I.P.A.T. Children's Personality Questionnaire.* Institute of Personality and Ability Testing, Illinois.

CATTELL, R. B. (1960 (*a*)). Predicting success in academic aspects of military training schools. *I.P.A.T. Information Bulletin No.* 4. Department of Psychology, University of Illinois.

EYZENCK, H. J. (1947). Student selection by psychological tests. *Brit. J. Educ. Psychol.*, 17, 20-33.

FURNEAUX, W. P. (1957). *Student Selection.* Report to the Imperial College of Science and Technology. Quoted in Lynn (1959).

HALLWORTH, H. J. (1961). Anxiety in secondary school children. *Brit. J. Educ. Psychol.*, 31, 281-291.

HOLMES, F. J. (1960). Predicting academic success on a general college curriculum. *I.P.A.T. Information Bulletin, No.* 4. Department of Psychology, University of Illinois.

LUNZER, E. A. (1960). Aggressive and withdrawing children in the normal school. *Brit. J. Educ. Psychol.*, 30, 119-123.

LYNN, R. (1959). Two personality characteristics related to academic achievement. *Brit. J. Educ. Psychol.*, 29, 213-217.

WARBURTON, F. W., BUTCHER, H. J., and FORREST, G. M. (1963). Predicting student performance in a university department of education. *Brit. J. Educ. Psychol.*, 33, 68-80.

(*Manuscript received 6th May, 1965*)

THE RELATIONSHIPS BETWEEN INTELLIGENCE, ANXIETY, NEUROTICISM AND EXTRAVERSION

BY *P. LEY, *M. S. SPELMAN, †ANN D. M. DAVIES AND ‡S. RILEY

SUMMARY. This investigation is concerned with testing the hypotheses that the regression of intelligence test scores on anxiety and neuroticism is curvilinear, and in the form of an 'inverted U,' and that extraversion is negatively correlated with intelligence. The results support neither prediction. The findings are discussed.

I.—INTRODUCTION.

INVESTIGATIONS of the relationships between anxiety, neuroticism, extraversion and intelligence test performance, have recently been reviewed by Lynn and Gordon (1961), Payne (1961), Child (1964). and Eysenck and White (1964).

The reviews suggest that anxiety and neuroticism can be conceptualized as drives. If this is so, consideration of the Yerkes-Dodson Law, leads to the expectation that there is a curvilinear relationship between measures of anxiety or neuroticism and intelligence—some moderate level of anxiety giving better performance than either low or high anxiety. Payne (1961) points out that many investigations have not considered the possibility of a curvilinear relationship. He argues that if such a relationship between anxiety and intelligence exists, results depend on the level of anxiety in the sample studied. Payne points out that differing results have, in fact, been found : Kerrick (1955) and Calvin, *et al.* (1955) found that intelligence was correlated negatively with scores on Taylor's Manifest Anxiety Scale (M.A.S.), while Schultz and Calvin (1955) and Mayner, *et al.* (1955) found no significant correlation between M.A.S. and intelligence.

Evidence of a curvilinear relationshp between anxiety and intelligence has been found by Sarason and Mandler (1952) and also by Lynn and Gordon (1961).

Eysenck and White (1964) have considered the relationship between extraversion and intelligence and argue that performance on many intelligence tests consists of attempting several very similar items in succession. Massed practice effects might be expected to appear under such circumstances and reactive inhibition build up. Extraverts, having a greater tendency to accumulate reactive inhibition, should show a fall-off in test performance before Introverts, and their test scores might be expected to be lower than those of Introverts in the absence of any true difference in intelligence between the groups.

The present study is designed to investigate two hypotheses : (*a*) that the regression of intelligence test scores on measures of anxiety will depart significantly from linearity, and that the shape of the regression line will be that of an 'inverted U,' and (*b*) that there will be negative correlations between tests of intelligence and measures of extraversion.

---

* Department of Psychological Medicine, University of Liverpool.

† Medical Research Council Unit for Research on Occupational Aspects of Ageing, University of Liverpool.

‡ Salford College of Advanced Technology.

II.—METHOD.

The sample consisted of 144 normal volunteers drawn from the Liverpool M.R.C. Panel (Heron, 1962). The subjects had a mean age of 46·45 years, S.D. 13·10 years. All subjects had previously completed the Progressive Matrices Test and the Mill Hill Synonyms Test, Form II Senior.

The tests administered in the present investigation were the Maudsley Personality Inventory (M.P.I.), the I.P.A.T. Anxiety Scale, the Taylor Manifest Anxiety Scale (M.A.S.) and the 16 P.F. Questionnaire Form C. Factor B of this last test is a measure of intelligence.

Table 1 shows the means and standard deviations of the subjects' scores on these intelligence, anxiety and extraversion tests.

TABLE 1

SUBJECTS SCORES ON PERSONALITY MEASURES AND INTELLIGENCE TESTS.

Test	Sample	
	Mean	S.D.
M.P.I. E Scale	22·09	9·89
M.P.I. N Scale	20·31	11·44
Cattell Introversion	50·55	9·44
Taylor, M.A.S.	10·47	7·27
I.P.A.T. Anxiety	6·18	1·97
Progressive Matrices	44·35	9·70
Mill Hill Vocabulary	32·41	6·18
Factor B	5·25	2·31

The interrelationship of these variables was examined by means of product moment correlations. In addition, correlation ratios were computed for the regressions of intelligence test scores on anxiety and neuroticism measures, in order to test the hypothesis of a significant departure from linearity.

III.—RESULTS.

(1) *Anxiety, Neuroticism and Intelligence.*

Table 2 shows the correlation coefficients (r), correlation ratios (eta), and F and P values for tests of linearity, between measures of neuroticism or anxiety and intelligence.

TABLE 2

CORRELATION COEFFICIENTS, CORRELATION RATIOS AND TESTS OF LINEARITY OF THE REGRESSIONS OF INTELLIGENCE ON ANXIETY AND NEUROTICISM.

	r	eta	F	P
Progressive Matrices *vs.* M.P.I. N Scale	—·357†	·409	·926	n.s.
Progressive Matrices *vs.* Cattell Anxiety Scale	—·228†	·262	·500	n.s.
Progressive Matrices *vs.* Taylor M.A.S.	—·253†	·469	5·479	< ·01
Mill Hill Vocabulary *vs.* M.P.I. N Scale	·167*	·283	1·090	n.s.
Mill Hill Vocabulary *vs.* Cattell Anxiety Scale	—·105	·154	·279	n.s.
Mill Hill Vocabulary *vs.* Taylor M.A.S.	—·210*	·413	4·198	< ·01
Cattell Factor B *vs.* M.P.I. N Scale	—·290†	·346	·877	n.s.
Cattell Factor B *vs.* Cattell Anxiety Scale	—·219†	·385	3·200	< ·01
Cattell Factor B *vs.* Taylor M.A.S.	—·245†	·376	2·880	< ·05

*—P=< ·05. †—P=< ·01.

It can be seen from Table 2 that four of the relationships depart significantly from linearity. All three intelligence measures have a significantly non-linear regression on the M.A.S., and Factor B also has a significantly non-linear regression on the I.P.A.T. Anxiety Scale.

Regression lines were plotted for these four regressions to see if they followed the predicted inverted ' U ' pattern. None did so, but all were of the same general form. As an example, Figure 1 shows the regression of Matrices scores on the M.A.S.

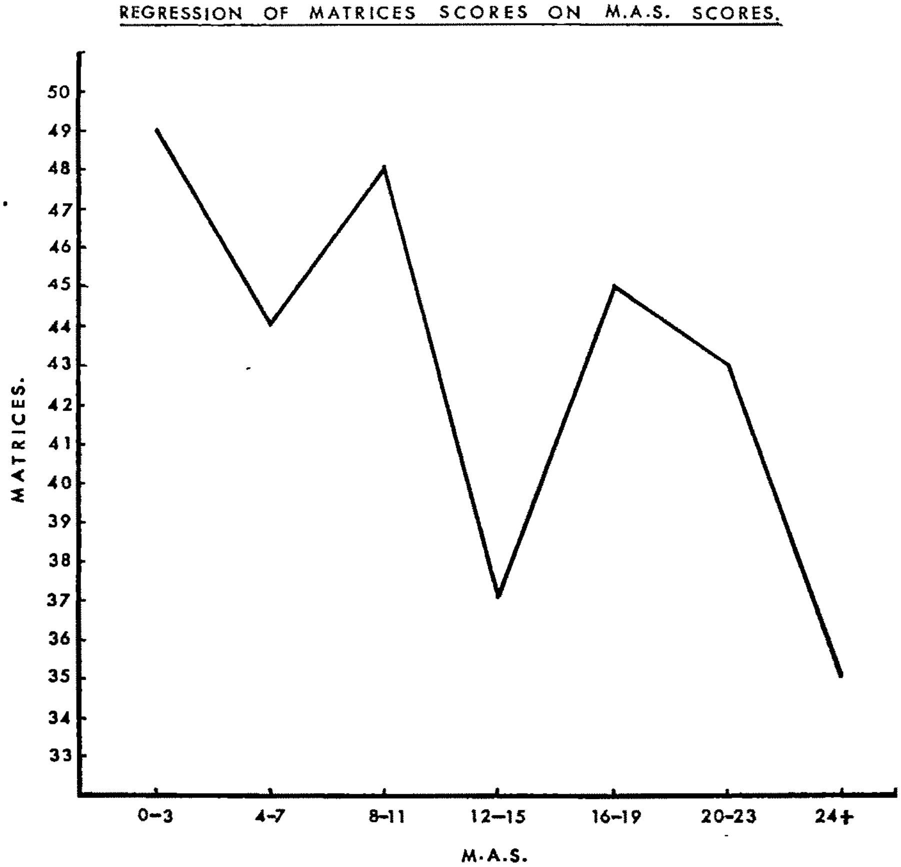


FIGURE 1

As can be seen, a moderate level of anxiety is associated with lower scores han most other levels. This is clearly contrary to the first hypothesis.

Cattell (I.P.A.T., 1957b) has argued that anxiety is made up of a number of components. As scores for these components were available, the regression of intelligence test scores on the components was investigated. The results are shown in Tables 3, 4 and 5.

TABLE 3

CORRELATIONS (r), CORRELATION RATIOS (eta), AND TESTS OF LINEARITY BETWEEN ANXIETY TESTS AND PROGRESSIVE MATRICES.

Test	r	eta	F	P
Factor C	—·137	·226	·764	n.s.
Factor L	—·222†	·268	·678	n.s.
Factor O	—·281†	·329	·890	n s.
Factor Q3	—·025	·123	·490	n.s.
Factor Q4	—·186*	·263	·827	n.s.

*—P=<·05. † P=<·01.

TABLE 4

CORRELATIONS (r), CORRELATION RATIOS (eta), AND TESTS OF LINEARITY BETWEEN ANXIETY TESTS AND MILL HILL VOCABULARY.

Test	r	eta	F	P
Factor C	·028	·102	·249	n.s.
Factor L	·094	·218	1·120	n s.
Factor O	·118	·204	·801	n.s.
Factor Q3	·051	·199	·107	n s.
Factor Q4	·098	·189	·905	n s

TABLE 5

CORRELATIONS (r), CORRELATION RATIOS (eta) AND TESTS OF LINEARITY BETWEEN ANXIETY TESTS AND CATTELL'S FACTOR B.

Test	r	eta	F	P
Factor C	—·180*	·259	1·030	n s.
Factor L	—·125	·272	1·720	n.s.
Factor O	—·311†	·350	·840	n.s.
Factor Q3	—·112	·182	·404	n.s.
Factor Q4	—·173*	·277	1·759	n.s.

* P=<·05. † P=<·01.

It can be seen that none of the regressions departed significantly from linearity.

(2) *Extraversion and Intelligence.*

The correlations between the two measures of extraversion and the three intelligence tests are shown in Table 6. (It should be noted that the scoring of the Cattell Second Order Extraversion Factor is such that high scores are associated with introversion.)

TABLE 6

CORRELATIONS BETWEEN EXTRAVERSION TESTS AND INTELLIGENCE TESTS.

Test	Progressive Matrices	Mill Hill Vocabulary	Factor B
M.P.I. E Scale	·072	—·027	+·171*
16 P.F. 2nd order Factor Introversion	—·104	·086	—·172*

* P=<·05

Only two of the correlations are significant and these are in the wrong direction.

As Cattell (I.P.A.T., 1957a) has argued that extraversion can be broken down into various components, the relationships between the intelligence test scores and these components were examined. The results are shown in Table 7.

TABLE 7

CORRELATION COEFFICIENTS BETWEEN EXTRAVERSION COMPONENTS AND THE THREE INTELLIGENCE TESTS.

Factors	Progressive Matrices	Mill Hill Vocabulary	Factor B
A (Cyclothymia)	+·112	+·021	+·154
F (Happy-go-lucky)	+·110	+·031	+·178*
H (Adventurous)	+·260†	+·244†	+·303†
M (Bohemian Introversion)	+·162	+·327†	+·012
Q1 (Radical)	+·123	+·135	+·105
Q2 (Self-sufficient)	+·076	+·018	—·030

* P=<·05. † P=<·01

The table shows that, with the exception of the correlation between the Mill Hill Vocabulary Test and Factor M (Bohemian Introversion), all the significant correlations are in the wrong direction, with extraverts doing better than introverts.

IV.—DISCUSSION.

The results obtained in the present investigation lend no support whatever to the hypotheses tested. Of the twenty-four regressions of intelligence test scores on anxiety measures, most did not depart significantly from linearity and where there were such departures there was no evidence of an ' inverted U ' relationship. Nor did we find the predicted negative correlations between extraversion and intelligence. Indeed, with one exception, all significant correlations between extraversion and intelligence were positive.

Comparison of the present study with those in which an 'inverted U' relationship has been found reveals that the present sample is much more heterogenous. For instance, Lynn and Gordon (1961) used undergraduates, presumably a group of limited range of age and intellect. If the relationship between anxiety and intelligence test performance varies with differing ages or levels of ability, then the heterogeneity of our sample may have masked such relationships. When achievement (rather than intelligence) was examined by Spielburger (1962), it was found that students of low intelligence earned low grades, irrespective of anxiety level, while in students of medium intelligence there was a negative relationship between anxiety and achievement. In highly intelligent subjects, anxiety facilitated performance. Furneaux (1962), found that anxiety is correlated positively with achievement in university students (a high intelligence group), while Child (1964), found anxiety to be negatively correlated with achievement in comprehensive school children (a less intelligent group).

To investigate the possibility that heterogeneity of intelligence was masking certain relationships in our data, the relationship of anxiety (M.A.S.) to intelligence (Progressive Matrices) was assessed at three different levels of intelligence. The high intelligence group consisted of those scoring 50 or more on the Matrices, the medium group of those scoring 40–49, and the low group, those scoring 39 or less. The mean intelligence test scores for each level of anxiety within the ability groups are shown in Table 8.

TABLE 8

RELATIONSHIP OF ANXIETY AND INTELLIGENCE AT DIFFERENT LEVELS OF INTELLIGENCE.

High Abiltiy			Medium Ability			Low Ability		
M.A.S. Scores	Mean Matrices Score	N	M.A.S. Scores	Mean Matrices Score	N	M.A.S. Scores	Mean Matrices Score	N
0—3	52·33	9	0—3	46·33	6	4—7	33·67	6
4—7	54·00	9	4—7	45·30	10	8—11	33·67	3
8—11	54·19	16	8—11	45·75	4	12—15	30·33	12
12—15	51·67	6	12—15	44·57	14	16—19	34·25	4
16—19	53·27	11	16—19	45·58	4	20—23	31·00	4
20—23	52·71	7	20—23	42·00	4	24—27	31·50	6
24—27	54·00	4	24—27	44·60	5			

There was no evidence of an 'inverted U' relationship at any level of ability. The relationship seeming closest to this was found in the high ability group. The correlation coefficient between M.A.S. and Matrices for this group was +·100 and the correlation ratio of Matrices on M.A.S. was ·355. A test of linearity yields an F value of 1·46 which is not significant. Thus, our failure to find an 'inverted U' is not due to heterogeneity of intelligence in our sample.

Our failure to find a negative correlation between extraversion and intelligence could be due to a variety of factors. Eysenck's prediction stresses the homogeneity of test items, and the length of the test. It is possible that the tests used were neither homogeneous enough nor long enough for reactive inhibition to build up to the point where it caused a decrement in test performance. It is hard to see that these objections apply to the Progressive Matrices, but they may possibly apply to our other tests.

Another possible explanation is that our extraverted subjects really were brighter than our introverted subjects, and that this factor masked the effects of a differential decrement in performance. Eysenck (1959) has shown that in extraverts, speed of producing correct solutions drops in the later stages of a test, and this suggests that it may be worth examining intelligence test scores in relation to personality at different stages of testing. Thus, if our extraverted subjects were brighter than our introverted subjects, and if their greater tendency to build up inhibition had so lowered their scores on a long test, that they appeared no brighter than the introverts, one would expect that at an earlier stage in the test there would be a higher correlation between extraversion and intelligence, than at the end of the test.

Scores were available for the Progressive Matrices test at 20 minutes and 40 minutes, and these were correlated with the M.P.I. E Scale. Neither of the correlations was significantly different from zero, nor did they differ significantly from each other.

V.—REFERENCES.

CALVIN, A. D., KOONS, P. B., BINGHAM, J. L., and FINK, M. M. (1955). A further investigation of the relationship between manifest anxiety and intelligence. *J. Cons. Psychol.*, 19, 280-282.

CHILD, D. (1964). The relationship between introversion and extraversion, neuroticism and performance in school examinations. *Brit. J. Educ. Psychol.*, 34, 187-196.

EYSENCK, H. J. (1959). Personality and problem solving. *Psychol. Rep.*, 5, 592.

EYSENCK, H. J., and WHITE, P. O. (1964). Personality and the measurement of intelligence. *Brit. J. Educ. Psychol.*, 34, 197-202.

FURNEAUX, W. D. (1962). The psychologist and the University. *Universities Quart.*, 17, 33-47.

HERON, A. (1962). The Liverpool functional age project: preliminary communication on population and methodology. *Experientia*, 18, 470.

INSTITUTE FOR PERSONALITY AND ABILITY TESTING (1957). *Handbook for the* 16 *P.F. Questionnaire*. Champaign, Illinois.

INSTITUTE FOR PERSONALITY AND ABILITY TESTING (1957). *Handbook for the I.P.A.T. Anxiety Scale*. Champaign, Illinois.

KERRICK, J. S. (1955). Some Correlates of the Manifest Anxiety. *J. Abnorm. Soc. Psychol.*, 63, 75-77.

LYNN, R., and GORDON, E. (1961). The relationship of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 194-203.

MAYNER, M. S., SERSEN, E., and TRESSALT, M. E. (1955). The Taylor manifest anxiety scale and intelligence. *J. Cons. Psychol.*, 19, 401-403.

PAYNE, R. W. (1961). Cognitive abnormalities in Eysenck, H. J., *Handbook of Abnormal Psychology*. London: Pitman.

SARASON, S. B., and MANDLER, G. (1952). Some correlates of test anxiety. *J. Abnorm. Soc. Psychol.*, 47, 810-817.

SCHULZ, R. E., and CALVIN, A. D. (1955). A failure to replicate the finding of a negative correlation between manifest anxiety and A.C.E. scores. *J. Cons. Psychol.*, 19, 223-224.

SPIELBURGER, C. D. (1962). Effects of manifest anxiety on the academic performance of college students. *Ment. Hyg.*, New York: 46, 420-426.

(Manuscript received 17th May, 1965)

EVALUATIONS OF FORCED-CHOICE AND LIKERT-TYPE TESTS OF MOTIVATION TO ACADEMIC ACHIEVEMENT

BY CLAUDE E. BUXTON

(*Yale University*)

SUMMARY. One forced-choice test of academic motivation and two Likert-type tests were developed and their validity assessed. One of the Likert tests appeared to possess little validity, for several possible reasons. But the forced-choice test and a Likert test constructed from the same materials (sentences) and administered to comparable groups under comparable conditions appeared to be about equally valid, as judged by significant relations with teacher ratings and stream level of the pupil.

I.—BACKGROUND OF THE PROBLEM.

Purposes of the research. The long-term purpose of these researches is to investigate how the behaviour of the teacher influences the motivation (attitude to attainment) of the pupil. To do such observational or experimental research in the classroom one must be able to secure a measure of the strength of motivation which is independent of the teacher's judgment of the motivation of any particular pupil. In the beginning, however, one can not do better than construct an objective measure of motivation by requiring that it produce results in accord with teachers' normal evaluations of pupils and with other criteria assumed to indicate the pupil's motivation directly or indirectly. I, therefore, report here on the construction of two forms of pencil-and-paper test of strength of academic motivation and record how they relate to teacher judgments and other criteria.

Theory underlying the tests. Everyone knows that a pupil's strength of motivation to do well in school work varies from class to class, teacher to teacher, or time to time. I shall not go into detail, for a theoretical background for understanding these variations is given by Atkinson (1964, esp. Ch. 10). Central to his theorizing are three concepts : strength of expectancy of success or failure, in attaining a goal ; degree of desirability or repulsiveness of the goal ; and strength of persisting motives which may negate or supplement one another. The tests discussed here are efforts to measure objectively an instance of the third of these, the persisting motive or need to achieve in school. If such a test can be developed, attention may then be turned to the way a pupil's answers to it, and his goals and expectancies, are influenced by the teacher's behaviour. At the present time, I am concerned to develop a research instrument, not a practical device.

II.—CONSTRUCTION AND EVALUATION OF A FORCED-CHOICE TEST.

The problem of social desirability. Any answer a pupil gives in a test asking personal questions may be influenced by what others might think or what he would like to believe. In the American literature, this is referred to as the social desirability factor. It might also be called the influence of the moral evaluation the pupil believes or wishes an answer would be given by others, and the terms will be treated as synonymous in this paper.

The forced-choice type of test is designed to reduce the influence of this factor, and although its use has been controversial, a recent review suggests that in a good many situations it is reasonably valid (Zavala, 1965). In one of its simplest forms this type of test presents two statements and the person taking the test is ' forced ' to choose the one more nearly true of him (or less untrue). The two statements are chosen jointly on two criteria: they should appear as nearly equal as possible in social desirability (or undesirability) to the person taking the test; and they should be as different as possible, in the judgment of experts, with respect to what they indicate about the strength of the trait which the whole test purports to measure. Consider the following pair of statements: " I think it is more important to have friends than to be top of the class." " I am determined to do the best work I can." These sentences were rated by 439 grammar school boys on a scale ranging from –5 (very undesirable thing to say or do) to +5 (very suitable). The first had an average rating of 2·80 (fairly good thing to say) and the second 2·34 (also a fairly good thing to say). Ratings of the strength of academic motivation revealed by a pupil who ticked either sentence as true of himself were made by twenty-four psychologists who worked in schools, also on a scale from –5 to +5. The average motivation rating of the first sentence was –1·09 and of the second 3·25. Two such sentences are paired and a pupil asked to tick one, as indicated above. It is then hoped that with average or typical moral evaluations of the statements, or the behaviour they imply, controlled (but certainly not completely for any specific pupil) the answers the pupil selects will truly reveal something of his motivation. Dissembling, whether intentional or not, should be reduced.

Note that pupil ratings were used to equate statements for social desirability, but psychologist ratings to select statements for differences in the strength of motivation revealed. This was done because in the preliminary study referred to above, the twenty-four psychologists and also fifty-one teachers were asked to rate a large number of sentences for both strength of academic motivation revealed and for social desirability as they thought this would be judged by pupils. For both groups, the ratings on both dimensions were almost exactly the same, on the average. That is, these professional groups, presumably having certain standards in mind for their pupils, judged the motivational value and the social desirability of the sentences (or the behaviour implied) to be almost identical. Yet the social desirability ratings by pupils did not correlate very closely with either set of ratings by either group of adults, suggesting that pupils interpret the sentences or the behaviour according to their own standards. Hence, the effort to reduce the influence of social desirability upon answers must be based upon pupil standards concerning what is desirable.

Preliminary test construction. On the basis of the study just mentioned, thirty-nine pairs of sentences having social desirability ratings that seemed relatively stable from one form-level to another, were combined into a test and administered to 904 boys in two grammar schools and 374 boys in a modern school. They were promised their answers would be confidential. The sentences had been edited to remove Americanisms.

It was necessary next to discover which of the thirty-nine items were answered in a way consistent with other criteria of strength of motivation. At first, I had expected to be able to use marks as a criterion, but had to give this up because the way in which marks are given varies so widely from class to class and stream to stream. However, it was possible to secure for all pupils a rating of strength of motivation by at least one teacher who knew each pupil well. (Strength of motivation was defined for the teacher as " *desire* to do well in school," *not* actual attainment.) These ratings were made on a scale from 1 (very weak motivation) to 5. If two teachers rated the pupils, ratings were averaged. It was also possible, of course, to learn the stream level of each boy, and although the placement of pupils in streams is a complex matter, motivation is one consideration. (A chi-square test showed teacher ratings of strength of motivation and stream level to be very significantly related.) Therefore, the relation of test answers to stream level was investigated except where stream had a special definition such as subject-matter interest or administrative necessity. Because the test was not intended to measure ability, the tendency of each item to relate to I.Q. scores was investigated. I also checked on the relation between answers to any item and the form level of the pupils, not because form level is a criterion of motivation, but rather one of the variables of interest in any study focused on the schools.

It will be remembered that within each of the thirty-nine pairs of sentences one had either a more positive or less negative motivation rating than the other. This was termed the correct sentence in the pair. The first statistical analysis of results was made by the chi-square test, frequency counts being made for the number of correct answers in, for example, the ascending categories of teacher ratings for strength of motivation. Nineteen pairs of sentences showed significant relationships between the answer chosen and teacher ratings as well as stream levels. These pairs were unrelated to I.Q. in the two grammar schools and barely and somewhat inconsistently in the modern school. Then the nineteen items were treated as a test (see Appendix A), and a score derived for each pupil by counting the number of pairs in which he ticked the correct sentence as true of himself. Analyses of variance then demonstrated that the test as a whole was related as expected to each of the criteria.

Cross-validation. Essential in evaluating the test is whether it will also relate significantly to the criteria when tried with a new sample of subjects, where there will possibly be different causes of dissembling or different social pressures at work. For this evaluation two boys' grammar schools supplied, respectively, 215 and 392 subjects. The analyses of their data are shown in Table 1. It is clear that the nineteen-item test is very significantly related to teacher ratings. It fails, however, to relate significantly to stream level in School E and is related at just better than the ·05 level to stream level in School D. No I.Q. data were available in School D, but the test is clearly not related to I.Q. in School E. With one aberrant mean for each school, form level relates in a curvilinear manner to test score, but the variations in size of mean are not great and the significance level modest.

In the test, the sentences rated as indicating stronger motivation are predominantly (but not exclusively) statements reflecting a sense of duty and responsiveness to authority, or statements chosen to avoid a sentence indicating a negative attitude toward duty or authority. (Special care was taken in the instructions to be sure pupils knew how to choose an answer when each seemed essentially undesirable. Depending upon the individual, this difficulty could arise anywhere in the test.) There are only a few items suggesting fear of failure, a factor which might have been expected to be more conspicuous, according to the writings of Atkinson (1964) and Sarnoff, *et al.* (1959).

Items discarded during the process of selection included a few comparable to those in Appendix A, but items such as the following also did not have close enough relations to the criteria to cause them to be selected : " I enjoy reading books that go beyond our required reading in certain subjects ; I like to surprise a teacher by the good quality of my work ; When I finally get a difficult sum right I feel triumphant ; I get excited when I come out well in a lesson."

TABLE 1

ANALYSES OF VARIANCE AMONG FORCED-CHOICE TEST MEANS FOR FOUR VARIABLES.

	Grammar School D			Grammar School E		
	M	S.D.	F, d.f., level	M	S.D.	F, d.f., level
Teacher Rating*						
2	8·9	4·6	5·29	7·2	4·0	24·9
3	9·5	3·7	3/211	9·5	4·1	3/386
4	10·7	4·0	< ·01	10·1	3·8	< ·01
5	12·4	2·7		12·9	3·3	
Stream						
1	8·5	4·2	3·36	8·8	4·0	·88
2	9·4	4·1	2/135	9·4	4·2	2/236
3	10·6	3·8	< ·05	9·9	4·1	—
Ability Level†						
1				9·8	4·2	·65
2				10·3	4·2	3/342
3				9·5	4·2	—
4				9·5	4·4	
Form						
1	11·6	4·1	2·37	10·2	4·5	2·64
2	9·5	3·9	5/209	10·5	4·7	5/386
3	8·5	4·2	< ·05	8·8	3·9	< ·05
4	10·2	3·7		9·2	4·2	
5	9·9	4·4		9·1	3·7	
6	10·8	3·0		10·7	4·0	

* In neither school was a teacher willing to assign a 1 (very low) motivation rating to a pupil; this was a violation of rating instructions, but reflected, probably, the spirit or morale level of the schools as well.

† Because there were few I.Qs. below 120, and the ceiling is arbitrarily set at 140+, all I.Qs. in the school were divided into quartile groupings.

Thus, one does *not* get a feeling of importance of intellectual excitement, discovery, exploration or competitiveness, as aspects of academic motivation.

Summary, forced-choice test. When validated on pupils from two additional schools this type of test continues to predict teacher ratings significantly. It does not predict stream level so satisfactorily, i.e., there is 'shrinkage' which would affect a validity coefficient if one could be computed for streams. The test is unrelated to I.Q. in the school where data were available, and it is related in an approximately curvilinear manner to form level. The relation to teacher ratings indicates further investigation of this form of test, to determine whether its validity justifies its use as an independent measure in research.

III.—CONSTRUCTION AND EVALUATION OF A LIKERT-TYPE TEST.

Much simpler to construct than a forced-choice test is the type of test originated by Likert (1932). It does not control for social desirability or dissembling at all, but may be useful in research situations where the pupil knows his answers are confidential. Hence, it is evaluated here, to determine how its validity compares with that of the forced-choice test.

Preliminary test construction. An example of this kind of test item is the following : "When I get a poor mark I try to find out what I did wrong : ....Very true of me ; ....Partly or usually true of me ; ....Cannot say, no feeling one way or the other ; ....Partly or usually *untrue* of me ; ....Very *untrue* of me."

From the population of sentences referred to earlier, I selected forty for preliminary construction of a Likert test. In the initial study of social desirability ratings their means were somewhat more variable across form levels than were the means for the forced-choice sentences. The test was administered to 550 boys in two modern schools. Scoring was determined by the ratings the experts had given the sentence for academic motivation. If they gave it a positive rating (as they would the illustrative sentence above), the answer " very true of me " was scored 5 and so on down to 1 for " very *untrue* of me." For a negative sentence, these numbers were reversed. Thus, the sentences could be analyzed by noting how the frequencies with which each answer occurred related to the criteria employed previously, namely, teacher rating, stream level, I.Q. (and, for information merely, form level). Chi-square tests indicated that seventeen of the forty items were related to the criteria as they should be. They were combined into a test (see Appendix B). The total scores were checked by analyses of variance, and it was ascertained that they related to the criteria as expected.

Since it has been shown (Festinger and Katz, 1953, p. 530) that to be unambiguous a Likert-type statement has to be strongly worded, a check was made on the academic motivation ratings assigned to these seventeen items by the experts referred to earlier. The average rating by the psychologists was 2.87, where the maximum would be 5, disregarding algebraic sign. The seventeen sentences are, therefore, mostly strong statements, in the judges' opinions.

Cross validation. For this purpose the seventeen-item test was administered to 285 boys in a grammar school and 114 boys in a technical school. Analysis of variance again was used to test the relations between test score and the various criteria, as shown in Table 2.

First, and in contrast with results with the forced-choice test, the relation between teacher rating and test score is insignificant in the technical school and reaches only the ·05 level in the modern school. Next, stream level is significantly related to test score, but the direction of the relation is opposite in the two schools, for unknown reasons. I.Q. is unrelated to test score in either school, as it presumably should be. On the other hand, form level is related to test score in a curvilinear manner in the modern school (as was the suggestion for the forced-choice test) but in a linear manner in the technical school.

Summary, Likert-type test. In contrast with the forced-choice test, a Likert test constructed by selecting items significantly related to certain criteria of motive strength in an initial set of schools shows quite a severe and partially confusing shrinkage of the relations between criteria and test when it is administered to new school populations. Three factors may be involved. First, the variation in answers across form levels reduces the possibility that these answers *could* relate significantly to teacher rating or stream level. Second, items selected because of significant relations with the criteria in a modern

TABLE 2

ANALYSES OF VARIANCE AMONG LIKERT TEST MEANS FOR FOUR VARIABLES.

	Technical School A			Modern School C		
	M	S.D.	F, d.f., level	M	S.D.	F, d.f., level
Teacher Rating						
	(2) 42·6*	7·3	2·07	(1) 52·8	13·0	3·07
	(3) 44·4	13·8	5/108	(2) 53·1	13·8	4/278
	(4) 49·4	13·8	—	(3) 52·7	11·7	< ·05
	(5) 54·4	12·0		(4) 56·2	11·9	
	(6) 53·1	11·6		(5) 61·0	11·1	
	(7) 52·5	9·7				
Stream						
1	49·0	11·2	6·70	58·6	13·8	3·80
2	57·7	9·8	1/40	52·6	11·5	2/265
3			< ·05	54·0	13·3	< ·05
Ability Level†						
	(3) 61·8	13·8	0·11	(1) 56·4	13·4	1·17
	(4) 61·0	11·2	2/99	(2) 56·1	11·0	4/269
	(5) 52·2	12·1	—	(3) 53·4	12·4	—
				(4) 53·5	12·7	
				(5) 51·2	11·8	
Form						
1	57·9	9·2	7·68	56·9	12·1	6·26
2	56·2	8·2	4/107	57·0	13·2	4/280
3	54·8	11·3	< ·01	54·5	12·1	< ·01
4	47·7	12·4		49·2	11·6	
5	44·2	10·6		60·8	9·5	

* Ratings of two teachers averaged and rounded, here only.
† Code 1 means I.Q. 70-80 ; 2, 80-90, and so on.

school are less appropriate for a grammar and technical school population Third, the sentences in the two tests were different, and as cross-validation was carried out in different types of schools, the content of the tests may have been a complicating factor.

IV.—COMPARISON OF FORCED-CHOICE AND LIKERT-TYPE TESTS WITH COMPARABLE CONTENT.

To control in part for all three of the factors just mentioned, I constructed a Likert test by arbitrarily selecting from the nineteen-item forced-choice test described in Part II, the nineteen sentences having the more extreme motivation rating (whether positive or negative in algebraic sign), and casting them into the Likert format. Data were gathered with this test in the same schools as for the cross-validation of the forced-choice test and at the same time, by giving every other pupil a Likert test and his neighbours a forced-choice test. Comparison of results with the two test forms is more readily justified with this procedure. (The items used for this Likert test can be found in Appendix A by noting, in each pair of sentences, the one with the larger academic motivation rating, whether positive or negative; scoring was direct or reverse as for the independently derived Likert test of Part III. Their mean rating was 2·75, again indicating that the selected statements were rather strongly and presumably unambiguously worded.)

Comparison of results. Table 1 showed that teacher ratings were significantly related to forced-choice score in both schools used for cross validation. In Table 3 we see that a Likert test constructed by the arbitrary method

TABLE 3

ANALYSES OF VARIANCE AMONG LIKERT-TEST MEANS FOR FOUR VARIABLES.
19-Item Test.

	Grammar School D			Grammar School E		
	M	S.D.	F, d.f., level	M	S.D.	F., d.f., level
Teacher Rating*						
2	61·2	11·3	14·14	59·6	12·7	26·10
3	66·9	10·8	3/211	66·6	10·6	3/387
4	72·7	10·3	<0·1	70·7	10·1	<0·1
5	76·0	7·1		75·5	10·4	
Stream						
1	63·0	12·1	13·24	63·5	12·7	6·85
2	68·2	9·2	2/136	66·7	12·1	2/238
3	73·5	9·0	< ·01	72·6	11·4	< ·01
Ability Level†						
1				65·5	13·5	1·44
2				66·3	11·6	3/333
3				69·0	11·0	
4				77·6	11·1	
Form						
1	70·6	12·6	·74	67·4	11·2	2·30
2	69·5	10·1	5/209	69·6	10·4	5/385
3	70·3	9·2	—	66·8	13·1	< ·05
4	67·5	12·3		64·5	13·0	
5	68·0	11·8		66·5	12·9	
6	71·4	10·6		70·4	10·4	

* In neither school was a teacher willing to assign a 1 (very low) motivation rating to a pupil; this was a violation of rating instruction, but reflected, probably, the spirit or morale level of the schools as well.

† Because there were few I.Qs. below 120, and the ceiling is arbitrarily set at 140+, all I.Qs. in a school were divided into quartile groupings.

described above also relates very significantly to teacher ratings, thus contrasting sharply with the results of the independently constructed Likert test of Part III. Table 1 also showed that the forced-choice test did not relate impressively to stream-level, whereas we see in Table 3 that the Likert test is related at a high level of significance. In the one school in which I.Qs. were available, they were unrelated to the Likert test, as was true for the forced-choice test. In one of the two schools the Likert test was barely significantly related to form level, whereas in the other it was not related, a result somewhat less clear than that obtained in cross-validation of the forced-choice test and considerably less clear than the significant relations reported for the independently constructed Likert test (Table 2).

Summary of comparisons. A Likert test in which the sentences are simply the nineteen 'strongest' sentences of the forced-choice test with which it is being compared shows more significant relations with teacher ratings and stream level than does the forced-choice test. Neither test was related to I.Q. and the curvilinear relation to form level previously is less clear but suggested in these results for both types of test. The Likert test, therefore, deserves further trial also.

The finding here that there is not a clearly curvilinear relation between test score and form level suggests that stability of social desirability ratings across forms is requisite to valid relations of test scores with teacher ratings and stream level.

V.—DISCUSSION AND CONCLUSIONS.

The data presented in Parts II and IV of this paper show that it is possible to construct either a forced-choice test or a Likert test of academic motivation which will relate significantly to important criteria in validations with new groups of subjects. This is not to say that correlation coefficients, if they could be computed, would be high, but examination of the trends of means presented in Tables 1 and 3 suggests that, especially for research purposes, the correlations are real. (This does not imply that the tests could safely be used for individual measurement.)

There is no certain explanation of the failures of the Likert test discussed in Part III. The rather complete reversal of findings with the Likert test when I used a different set of sentences, more stable in their social desirability ratings among different form levels, suggests that to measure presisting motives to do well in school it is most valid to select those sentences implying attitudes or behaviours which have stablized by the time the pupil reaches secondary school.

Before reaching the conclusion that the two types of test *can* be equally meritorious other checks must be made to determine the effects of the variable of social desirability of an answer. Especially important are such questions as, what difference does it make if the pupil thinks someone in authority over him in the school might see his answers and, to what extent are his answers influenced by what other pupils might think ? Work upon such questions should also permit improvement of the preliminary tests developed here.

ACKNOWLEDGEMENTS.—My special thanks go to Professor P. E. Vernon and Dr. W. H. King of the University of London Institute of Education. In preliminary and later phases of the work, I had the co-operation of these schools : Sutton County School for Boys ; Queen's County Secondary Boys, Wimbledon ; Rayleigh County Secondary School ; Sweyne School (Grammar-Technical) ; Freyerns Grammar School ; Strode's School, East Ham Grammar School ;

Holbrook County Secondary School; Plaistow County Grammar School; Harold County Secondary School; Lister Technical School. Dr. Helen Culler and Mr. Frederick G. S. Turner assisted in data collection, and Miss Sarla Merchant and Messrs. Daniel D. Wheeler and Kurt W. Fischer in data analysis. Financial assistance has been given by the National Institute of Mental Health, small grant No. M-5963, and by the Carnegie Corporation and the Kettering Foundation. I am exceedingly grateful for all this aid.

VI.—REFERENCES.

ATKINSON, J. W. (1964). *An Introduction to Motiviation.* Princeton, N.J.: D. Van Nostrand Co.

FESTINGER, L., and KATZ, E. (Eds.) (1953). *Research Methods in the Behavioral Sciences.* New York: Dryden Press, Inc.

LIKERT, R. (1932). A technique for the measurement of attitudes. *Arch. Psychol.*, No. 140.

SARNOFF, I., SARASON, S. B., LIGHTHALL, F. F., and DAVIDSON, K. S. (1959). Test anxiety and the eleven-plus examinations. *Brit. J. Educ. Psychol.*, 29, 9-16.

ZAVALA, A. (1965). Development of the forced-choice rating scale technique. *Psychol. Bull.*, 63, 117-124.

VII.—APPENDIX A.

A.M.*	S.D.*		The 19-Pair Forced-Choice Test.
2·69	–0·74	1a.	I don't mind working at break, if there is work to be done.
–1·92	–0·62	1b.	I usually don't understand what my teacher is criticizing in my work.
–2·58	–2·87	2a.	I don't care much about my marks.
2·09	–2·75	2b.	I don't mind not having sport, if there is work to be done.
–2·37	1·34	3a.	Sports are more important to me than is school work.
–0·80	–1·65	3b.	Before a test I often lie awake part of the night worrying.
–1·33	0·13	4a.	It doesn't disturb *me* much to be told off by a teacher.
2·14	0·04	4b.	Whenever I catch myself being lazy I do something about it straightaway.
3·96	2·41	5a.	When I find that I don't understand something I try very hard to master it.
1·79	2·33	5b.	If I get a low mark I feel rather discouraged, but tell myself I must do better next time.
–3·56	–0·41	6a.	I often find my lessons so difficult I just don't want to think about them.
–1·00	–0·48	6b.	When I try too hard in my studies I become nervous and afraid I'll make a fool of myself.
–1·67	1·82	7a.	I think it more important to have a bit of fun than to work hard all the time in school.
3·33	1·69	7b.	I set standards for myself and then go out to achieve them.
1·75	–1·07	8a.	I like to have my friends be proud of my work.
–1·81	–1·11	8b.	I don't profit much from a teacher's criticisms of my work.
0·90	–1·62	9a.	I find it difficult to eat on the day I am to have an important test.
–2·43	–1·43	9b.	Much of what I learn in school makes no difference to me.
1·00	0·25	10a.	When someone beats me by a long way I don't much care, but if he beats me just a little I am very disappointed.
2·25	0·33	10b.	I don't like to waste time.
1·33	–1·22	11a.	I don't fool around in class.
–2·71	–0·87	11b.	I find it's no use trying to do things that are too hard for me in school.
2·06	0·30	12a.	I don't feel right if I let down in my work.
–0·58	0·33	12b.	There are times when I feel as though I shall never get anywhere in school.

1·56	0·79	13a.	My teachers have a right to expect me to do my best.
–3·08	0·65	13b.	I shall be happiest when I am old enough to leave school.
2·11	–0·46	14a.	I don't wait for the teacher to tell me to get on with my work.
–0·96	–0·06	14b.	I think teachers and parents make too much of neatness and tidiness in school work, for it isn't that important.
–1·09	2·80	15a.	I think it is more important to have friends than to be top of the class.
3·25	2·34	15b.	I am determined to do the best work I can.
2·88	–0·90	16a.	I like keeping busy with my work.
–2·65	–0·74	16b.	If a teacher doesn't like my work it doesn't particularly disturb me.
2·13	–·28	17a.	I pay strict attention to what the teacher says so I won't miss anything.
–0·64	–0·13	17b.	Before a test I feel irritable or jumpy.
–1·88	–0·64	18a.	I just sit quietly without raising my hand, so the teacher won't call on me to answer questions in class.
–3·40	–0·66	18b.	My parents are always at me to do better in school, and this this makes me feel like doing just the opposite.
–2·81	–2·10	19a.	I don't care whether I pay attention to the teacher.
–0·06	–2·02	19b.	I just can't keep from worrying about my schoolwork.

* A.M.=mean academic motivation rating; S.D.=mean social desirability rating.

VIII.—APPENDIX B.

The 17-sentence Likert-type Test*

A.M.	S.D.		
2·89	0·25	1.	I feel ashamed when I get a poor mark.
1·56	–1·29	2.	I am quiet in the classroom and get on with my work.
2·25	–0·08	3.	I try hard to please my teacher with my work.
–2·80	–2·43	4.	I don't mind copying the work of someone else if I haven't done it for myself.
2·96	1·08	5.	I like to complete every assignment given to me.
3·71	–0·59	6.	Each new lesson makes me eager to get on with it.
3·89	1·58	7.	Getting to university is an important aim in life.
3.43	–3·13	8.	I hate to miss school because I enjoy it so much.
–2·44	1·30	9.	I think most homework is a bore.
2·78	–2·58	10.	I enjoy taking tests.
–2·42	–0·03	11.	I hate maths.
–2·92	–0·23	12.	Often I am just *not* in the mood for work, and then I don't do it.
–4·14	0·05	13.	I don't like school.
3·92	1·44	14.	I enjoy reading books that go beyond our required reading in certain subjects.
2·69	–0·74	15.	I don't mind working at break, if there is work to be done.
1·33	–1·22	16.	I don't fool around in class.
2·71	–1·66	17.	I like to have homework, because in that way I am sure I learn more.

* The available answers were always those of the example in the text.

(*Manuscript received* 20th May, 1965)

A COMPARATIVE STUDY OF VALUE JUDGMENTS OF ADOLESCENTS

By H. J. HALLWORTH and G. WAITE
(*Department of Education, University of Birmingham*)

SUMMARY. 155 boys and 128 girls in the fourth year of secondary school judged each of thirty-six concepts on scales taken from Osgood's three dimensions. Preliminary analysis confirmed the factorial composition of the scales. For each subject one evaluative score was obtained on every concept and the scores were intercorrelated and factor analysed for boys and girls separately. Five principal components were extracted in each case, accounting for 57 per cent. of the variance. The axes were rotated successively to the Varimax and Promax criteria. The two sets of Promax factors are discussed in relation to the attitudes of adolescent boys and girls to success and ambition, school and authority, security, the opposite sex, parents and self. For each sex group two second-order factors were obtained from the correlations between Promax factors. In each case, the larger was a general factor. For boys, the remaining second-order factor was related to school and authority; for girls, it was a factor of reward, money and youth. The investigation demonstrates a procedure for determining the affective meaning of concepts important in the lives of adolesccent boys and girls.

I.—INTRODUCTION.

THE 'semantic differential' is a technique developed for the measurement of connotative meaning by Osgood (1957). It consists in the evaluation of verbal concepts against a set of bi-polar scales, each defined by a pair of adjectives. In evaluating each concept, the subject places a mark in one of the intervals on the scale according to the degree of relationship perceived by him between the concept and one or the other of the adjectives defining the scale, thus:

BOY Active : : x : : : : : : Passive

Factor analytic procedures have shown that when a number of scales are used in this way, they commonly fall into three major groups, representing the main dimensions of meaning: Evaluation, Potency and Activity. Scales defining these three 'connotative' dimensions can, therefore, be used to measure individual or group differences in the perceived meanings of concepts, as well as differences between the concepts themselves. It has also been shown that scales tend to change their factorial identity as the concepts to which they are applied vary, some scales being more stable than others. This, therefore, necessitates that scales be selected and appraised for factorial identity for any particular concept or group of concepts being judged.

According to Osgood, the evaluative dimension is identifiable as the attitudinal component of meaning, and scales defining this dimension may be used to obtain measures of attitude towards concepts. Further, they may be used in investigations into similarities of attitudes towards groups of related concepts.

The authors have reported such an investigation into the dimensions of attitude among adolescent girls (Hallworth and Waite, 1963). The present article compares dimensions of attitude found among groups of adolescents of either sex.

II.—THE INVESTIGATION.

(*a*) *The Semantic Differential.*

The form of the semantic differential used in the investigation consisted of thirteen seven-point scales, as shown in the diagram.

E 7	successful	: : : : : : : :	unsuccessful	1
	still	: : : : : : : :	moving	
	slow	: : : : : : : :	fast	
E 7	beautiful	: : : : : : : :	ugly	1
E 1	cruel	: : : : : : : :	kind	
	difficult	: : : : : : : :	easy	
E 7	clean	: : : : : : : :	dirty	1
	manly	: : : : : : : :	womanly	
E 7	wise	: : : : : : : :	foolish	1
E 1	dishonest	: : : : : : : :	honest	7
	hard	: : : : : : : :	soft	
	light	: : : : : : : :	heavy	
E 7	good	: : : : : : : :	bad	1

Seven scales were chosen which Osgood has found to lie on the evaluative dimension, and these are marked ' E ' in the diagram above. They are also numbered to show the direction of scoring, a high score indicating high evaluation. Of the remaining scales, three were chosen from the activity dimension and three from the potency dimension. They served to divert attention from the purpose of the test.

These scales were applied to forty-eight concepts, of which a full list is given elsewhere (Hallworth and Waite, 1963). The concepts were chosen, subjectively, from areas which appeared important in the life of adolescents.

(*b*) *Subjects.*

The subjects were 155 boys and 128 girls, aged 14+. They were all fourth-year pupils in three neighbouring schools on the edge of an urban area of the Midlands : one boys' secondary modern, one girls' secondary modern, and one co-educational grammar school. The semantic differential was answered in a single one-hour session which gave ample time for all subjects. Instructions for administration are given in the article already referred to.

III.—STATISTICAL ANALYSIS.

In performing the statistical analysis, use was made of an electronic computer. A preliminary analysis confirmed that the seven evaluative scales had high loadings on an evaluative factor when the subjects were English adolescents.

Using these scales, an evaluative score was obtained for each subject on each of the forty-eight concepts. For all the girls, a correlation matrix was obtained for the forty-eight scores, five principal components were extracted and the axes were rotated to the Varimax criterion. The resultant factors have already been described (Hallworth and Waite, 1963). Loadings on three duplicated concepts, and on three concepts included as ' opposites,' gave a measure of substantiation to the validity of these factors. The three duplicates were omitted from further analyses. Certain abstract concepts appeared to contribute little of psychological signficance since they were grouped together on one factor, apparently indicating the subjects' inability to make any distinction between them ; these also were omitted from further analyses. The remaining thirty-six concepts are listed in Table 1.

For every subject there was now an evaluative score on each of thirty-six concepts. Boys and girls were regarded as separate samples. For each a correlation matrix was computed for the thirty-six evaluative scores ; five principal components were extracted, accounting for 57 per cent. of the variance in each case ; and the axes were rotated orthogonally to the Varimax criterion. The Varimax factors were than rotated by the Promax method (Hendrickson and White, 1964) to obtain loadings on five oblique factors.* For each of the two samples of subjects, comparison of the two sets of factors revealed little substantial difference. In Table 1, therefore, only the Promax factors are given. The two matrices of correlations between these factors are given in Table 2.

Two second-order factors were obtained from each matrix of correlations between Promax factors by extracing two principal components and rotating these to the Varimax criterion. Loadings of the thirty-six variables on each of the second-order factors are given in Table 1.

IV.—RESULTS : DIMENSIONS OF BOYS' ATTITUDES.

(*a*) *Factor I* : *Success, Ambition and Reward.*

The first Promax factor from the boys' analysis has highest loadings as follows :

36. success	·81	32. reward	·66	12. work	·48
30. games	·72	33. study	·62	21. future	·47
26. cleanliness ...	·70	34. trust	·56	14. play	·46
13. ambition	·67	29. mother	·51		

This is a factor of success, ambition, reward and, in some degree, the future. Associated with these concepts are games, study and work, also trust and mother. The presence of ' cleanliness ' suggests a puritanical view of life. The factor represents the most important group of concepts obtained for boys.

(*b*) *Factor II* : *Teachers, School and Authority.*

High loadings on the second factor are :

11. teachers	·82	35. punishment..	·66	9. religion......	·46
31. policeman	·76	20. rules	·58	33. study	·42
5. school	·73	8. books	·52		

This represents an image of teachers, school and authority generally, which is not altogether flattering to our educational system. The boys' attitudes to teachers, school, books and study are related to their attitudes to policemen, punishment and rules. There are not, on this factor, any substantial loadings for such concepts as ' kindness,' ' home ' and ' parents.' All concepts with high loadings, except ' religion,' have low mean evaluative scores for boys.

---

* All computer programs used were written in Mercury Autocode by H. J. Hallworth, assisted by Ann Brebner. They form part of a system of programs which will become available from the Department of Education, University of Birmingham. The Varimax program used Kaiser's loose limit, no rotations being effected which were less than 1°. The Promax program raised loadings to the fourth power.

TABLE 1

LOADINGS OF 36 VARIABLES ON 5 PROMAX FACTORS AND 2 SECOND-ORDER FACTORS FOR BOYS AND GIRLS.

Variables	Boys' Factors					Girls' Factors					Second-Order Factors			
											Boys		Girls	
	I	II	III	IV	V	I	II	III	IV	V	I	II	I	II
1. Honesty..	–03	21	73	–14	17	–19	14	08	84	00	55	42	53	24
2. Home ...	05	05	67	–20	31	49	16	–07	35	–18	58	31	63	08
3. Dress	03	–05	62	22	08	50	06	–11	42	03	69	18	61	29
4. Kindness .	20	07	77	–12	–09	13	16	–08	65	14	54	37	57	39
5. School ..	–12	73	31	–06	18	04	–14	72	42	–17	33	74	73	–04
6. Woman ..	–20	–07	30	65	10	73	–02	–02	06	03	65	–04	58	22
7. Pleasure .	24	–06	54	29	–08	67	–26	–06	28	23	69	23	49	43
8. Books ...	–22	52	48	19	–10	06	–27	61	17	17	38	49	46	20
9. Religion..	21	46	31	–03	–18	09	27	17	38	18	26	60	67	36
10. Freedom	32	–21	47	16	–11	42	–37	06	11	40	55	10	23	46
11. Teachers	–14	82	–04	07	23	05	–08	76	25	–13	21	73	71	–04
12. Work	48	20	26	11	–19	–13	19	66	06	05	46	49	63	09
13. Ambition	67	17	01	01	–04	53	–03	–07	01	40	40	53	41	51
14. Play	46	05	14	17	07	51	10	13	–07	20	55	36	58	32
15. Boy	18	–21	13	69	10	–01	–11	70	–30	37	76	–01	36	26
16. Money ...	12	–04	–21	68	01	10	–15	17	13	54	43	01	28	56
17. Myself ..	02	01	16	13	51	49	17	13	–03	00	57	18	61	15
18. Sister	–24	21	06	09	50	25	34	26	–17	05	35	19	57	12
19. Father ...	24	–03	06	–04	67	00	64	16	02	–02	60	26	62	10
20. Rules	16	58	16	–08	16	–01	14	50	37	–09	33	70	71	07
21. Future ...	47	05	–05	35	01	68	00	14	–06	–06	52	31	60	09
22. Men	–02	09	–10	73	10	10	09	74	–32	08	56	11	57	04
23. Friend ...	10	13	08	42	–08	39	59	–08	–07	00	40	20	66	18
24. Girl	–13	04	18	44	40	55	34	–11	–12	18	67	13	56	32
25. Adults ...	15	31	–24	26	38	18	34	45	–03	10	40	40	76	20
26. Cleanliness	70	–15	13	–06	19	–16	76	–26	26	16	56	32	44	32
27. Youth ...	25	19	–13	30	32	00	19	01	–11	56	52	36	19	54
28. Neighbours	–02	25	–13	26	36	70	18	09	–31	–19	37	27	54	–07
29. Mother ..	51	–08	–07	–06	61	17	74	–07	05	–02	59	30	68	17
30. Games ...	72	–06	–11	05	11	68	12	04	12	–33	44	34	66	–09
31. Policeman	09	76	–08	–15	20	–07	76	13	18	–12	12	78	72	05
32. Reward ..	66	–15	25	05	–10	–24	18	32	06	63	51	27	36	61
33. Study ...	62	42	10	–21	–09	–04	21	43	38	04	26	73	71	19
34. Trust	56	13	06	–16	28	26	53	–15	19	11	45	49	63	32
35. Punishment	13	66	–32	14	–49	22	–02	37	–02	–57	–29	51	34	–49
36. Success ..	81	01	00	08	00	02	34	08	22	51	52	46	56	62

TABLE 2

TWO MATRICES OF CORRELATIONS BETWEEN PROMAX FACTORS.

	I	II	III	IV	V
I	1·0	·43	·50	·39	·38
II	·54	1·0	·33	·28	·29
III	·52	·48	1·0	·39	·40
IV	·45	·37	·35	1·0	·30
V	·35	·32	·26	·29	1·0

Correlations for boys' analysis are above the diagonal.
Correlations for girls' analysis are below the diagonal.

(*c*) *Factor III* : *Virtue, Home and Pleasure.*

The third Promax factor for boys has the following high loadings :

4. kindness	·77	2. home	·67	7. pleasure	·54
1. honesty	·73	3. dress	·62	8. books	·48
				10. freedom	·47

This factor may be labelled ' virtue, home and pleasure,' and we may note that it has a moderate loading on ' books,' and lower loadings on ' school ' and ' reward,' but no significant loading on ' teachers.' ' Punishment ' has a negative loading.

(*d*) *Factor IV* : *Money and the Future.*

High loadings on the fourth factor are :

22. men	·73	16. money	·68	24. girl	·44
15. boy	·69	6. woman	·65	23. friend	·42
				21. future	·35

The factor appears to be concerned with the wish to grow up, especially in hetero-sexual relations, and to have the money necessary for this.

(*e*) *Factor V* : *Parents and Self.*

High loadings on the fifth factor are :

19. father	·67	18. sister	·50	25. adults	·38
29. mother	·61	35. punishment	−·49	28. neighbours	·36
17. myself	·51	24. girl	·40		

This factor apparently represents self in relation to the rest of the family, especially parents. ' Punishment ' has a substantial negative loading, but ' home ' has a somewhat lower loading of ·31.

V.—RESULTS : DIMENSIONS OF GIRLS' ATTITUDES.

(*a*) *Factor I* : *The Feminine Image.*

The first Promax factor for girls has high loadings as follows :

6. woman	·73	7. pleasure	·67	3. dress	·50
25. neighbours	·70	24. girl	·55	17. myself	·49
21. future	·68	13. ambition	·53	2. home	·49
30. games	·68	14. play	·51	10. freedom	·42

These concepts are apparently related to pleasure, ambition and the girl's future as a woman. They form, in effect, a feminine image.

(*b*) *Factor II* : *Security.*

High loadings on Promax Factor II are :

31. policeman	·76	29. mother	·74	23. friend	·59
26. cleanliness	·76	19. father	·64	34. trust	·53

This group of factors is apparently related to security. It comprises those personalities who may be trusted, and notably does not include ' teachers,' although ' sister ' and ' girl ' have moderate loadings.

(*c*) *Factor III* : *The Image of School and Men.*

High loadings on Promax Factor III are :

11. teachers	·76	15. boy	·70	20. rules	·50
22. men	·74	12. work	·66	25. adults	·45
5. school	·72	8. books	·61	33. study	·43

This appears to be, in the first instance, an image of teachers, school, work and study. With these, however, are associated the concepts of ‘ men ’ and ‘ boy.’ The data suggest no reason for this association. It should be remarked, however, that the next highest loading is on ‘ punishment.’

(*d*) *Factor IV* : *Honesty and Kindness.*

The only two concepts with substantial loadings on this factor are: 1. ‘ honesty ’ (·84) and 4. ‘ kindness ’ (·65). Other concepts such as ‘ dress,’ ‘ school,’ ‘ religion,’ ‘ study,’ and ‘ rules ’ have somewhat lower loadings.

(*e*) *Factor V* : *Reward, Youth and Money.*

Concepts with high loadings on Promax Factor V are :

32. reward	·63	16. money	·54	10. freedom	·40
35. punishment	·57	36. success	·51	15. boy	·37
27. youth	·56	13. ambition	·40		

This factor is apparently the equivalent of Factor I in the boys’ analysis. In the girls’ analysis, success, ambition and reward have fewer associations, and this group of concepts together accounts for less of the total variance. It is to be noted that ‘ games ’ has a negative loading of –·33, and that whilst ‘ boy ’ appears in this factor, ‘ girl ’ does not appear on the factor in the boys’ analysis.

VI.—DISCUSSION.

The similarities and differences between the factors obtained from the two analyses make good sense psychologically. For the boys, the most important cluster of concepts centres around success, ambition, reward and the future, and includes both games and study. For the girls, the principal cluster is concerned with a feminine image and the future. Girls are apparently more concerned with themselves as women ; games are prominent in this cluster, study is not ; dress and home are important.

In both analyses there is a cluster of concepts concerned with teachers, school and rules, books and study. For the boys, policemen are also associated with this group of concepts, as also are punishment and religion. With the exception of the last, all of these concepts also receive low mean evaluative scores for the boys. In effect, it appears that education does not have pleasant associations. In the case of the girls, however, other concepts on this dimension are ‘ men,’ ‘ boy,’ ‘ work ’ and ‘ adult.’ Education appears to be associated with work, and is something of a male province ; and again, the concepts all have a low mean evaluative score and are associated with punishment.

For boys, the factor of teachers, school and authority is second in importance. For girls, however, the second factor is concerned with security, represented by such concepts as ‘ mother,’ ‘ father,’ ‘ friend ’ and ‘ trust.’ Also on this factor is the concept of ‘ policeman,’ apparently represented in the role of guardian in contrast to his appearance, in the second factor for boys, as an authority figure. From the boys’ analysis, one factor appears comparable to the girls' 'security' factor. This extracts relatively little variance, however, and has been designated merely 'parents and self.'

The two virtues of honesty and kindess form the centre of another cluster of concepts in both analyses. For the girls, they have relatively unimportant associations. For the boys, they are associated with home, dress, pleasure, books and freedom.

The remaining group of concepts for boys, labelled 'money and the future,' is assumed to represent an attitude towards growing up, with which are associated money and the opposite sex. The last group of concepts for girls is similar in that it associates money with the opposite sex. In other respects, however, it may be compared to the first group of concepts for boys: in both reward, success and ambition are prominent. However, these have more associations for boys, and are related to games, study, mother and work. For girls, they are related to youth, money and freedom, and have a negative association with games.

The concept 'cleanliness' is of special interest. For boys, its highest loading is on the factor of success, ambition and reward; for girls, on the factor of security. In each case it is associated with a somewhat lower loading on 'mother' and 'trust.' It is apparently to be understood as representing clean-living, and the pursuit of things associated with this. In effect, such things are, for a boy, ambition and work, and for a girl, secure personal relationships.

VII.—SECOND-ORDER FACTORS.

Certain of the similarities and differences between the boys' and girls' attitudinal associations may be appropriately illustrated by a consideration of the two second-order factors obtained in each analysis. For the boys, these factors accounted for 64·53 per cent. of the variance of the Promax factor matrix; for the girls, they accounted for 67·44 per cent. Loadings of the thirty-six variables on the four second-order factors are given in Table 1.

In each analysis, the first of these factors is almost a general factor; for the boys, there are only five concepts with loadings of less than ·3; for the girls, there are only three such concepts. The highest loadings are different in the two cases, however. For the boys they are, in order, 'boy,' 'dress,' 'pleasure,' 'girl,' 'woman,' 'father,' 'mother,' 'home' and 'myself.' The prominent concepts here concern pleasure, girls, parents and home. For the girls, the highest loadings are, in order, 'adults,' 'school,' 'policeman,' 'teacher,' 'rules,' 'study,' 'mother' and 'religion.' This group of concepts centres around authority figures, but is by no means confined to these, and the girls' factor accounts for more of the total variance than does the boys'.

The remaining second-order factor is more sharply contrasted in the two analyses. For the boys, concepts with high loadings are 'policeman,' 'school,' 'teacher,' 'study,' 'rules' and 'religion'; 'punishment' is somewhat lower. This is obviously a factor of school and authority, with 'books' and 'ambition' having moderate loadings. It is comparable to the boys' Promax Factor II. For the girls, concepts with high loadings are 'reward,' 'money,' 'youth,' 'freedom,' and 'punishment,' the last loading being negative. The factor is comparable to the girls' Promax Factor V.

It is to be noted that 'myself' has a high loading for second-order Factor I for both sexes. Certain concepts have a considerably higher loading on one factor than on the other. Of these, 'school,' 'religion,' 'teacher,' 'rules,' 'policeman' and 'study' have a high loading on Factor II for boys and on Factor I for girls. In contrast, 'freedom,' 'money' and 'youth' have a high loading on Factor I for boys and on Factor II for girls. 'Punishment' has a positive loading on Factor II for boys, and a negative loading on Factor I; for girls, the opposite is the case.

VIII.—CONCLUSION.

The results of the two analyses have an obvious overall similarity. Differences between them are in general of the kind that would be anticipated. When we examine the Promax factors, boys have more attitudinal associations with ambition and study, girls with self and home; boys have a cluster of concepts identified with authority, girls with security. The girls' association of teachers and school with men and boys is less expected.

When we consider the second-order factors, we see that in both cases a large general factor is obtained, identifying a cluster of concepts with which both sexes associate themselves. In the case of the boys, another cluster, independent of the general factor, centres around school and authority. In the case of the girls, a similar cluster centres around reward, money and youth, whilst the concepts related to school and authority form part of the general cluster. It is apparent that there are sex differences in the attitudinal associations of these two groups of concepts.

It is not claimed that these two analyses provide a description of adolescent perceptions of the world which has not previously been available. It is, however, suggested that most other such descriptions have been subjectively based. The research described in this article gives an empirically based description of part of the adolescent world in quantitative terms. As such, it demonstrates a technique for the study of this type of phenomena.

It is now practicable to map an important aspect of subjects' perceptions or, in Osgood's terms, connotative meanings, over a wide area. Further, it appears practicable to assess such perceptions for groups of subjects contrasted by age, sex, social class, or by any attributes which appear relevant to the understanding of social problems.

IX.—REFERENCES.

HALLWORTH, H. J., and WAITE, G. (1963). A factorial study of value judgments among adolescent girls. *Brit. J. Math. Statist. Psychol.*, 16, 37-46.

HENDRICKSON, A. E., and WHITE, P. O. (1964). Promax: a quick method for rotation to oblique simple structure. *Brit. J. Math. Statist. Psychol.*, 17, 65-70.

KAISER, H. F. (1959). Computer program for Varimax rotation in factor analysis. *Educ. Psychol. Meas.*, 19, 413-420.

OSGOOD, C. E. (1962). Studies on the generality of affective meaning systems. *Amer. Psychol.*, 17, 10-28.

OSGOOD, C. E., SUCI, G. J., and TANNENBAUM, P. M. (1957). *The Measurement of Meaning*. Urbana: University of Illinois Press.

(*Manuscript received 1st June*, 1965)

AN EDUCATIONAL GUIDANCE PROGRAMME FOR THE PRIMARY SCHOOL

By NORMAN FRANCE
(*Institute of Education, McGill University*)

AND STEPHEN WISEMAN
(*School of Education, University of Manchester*)

SUMMARY. The France-Wiseman Educational Guidance Programme consists of three 32-page work-books which contain a variety of tasks for children to complete at their own pace as part of their normal class-work over a period of about two weeks. The three work-books are for use by first, second and third-year juniors, respectively. The separate tests were tried out in a variety of schools throughout the United Kingdom in the Spring of 1964, and were then standardised simultaneously in November, 1964, on a random sample of Warwickshire and Solihull primary schools. The scores were analysed first to provide tables of standard scores, and then to investigate the achievement of :

(*a*) children born in different months of the year,
(*b*) boys and girls,
(*c*) children from different areas of the county,
(*d*) children from different family backgrounds, and
(*e*) all children on the different sub-tests.

The results showed that untimed tests of this nature, even when of short duration, can produce a satisfactory discrimination over most of the ability range. (*a*) The effect of month of birth decreases for the older age-groups. (*b*) In all cases the spread of scores for girls is less than that for boys. Girls generally achieve higher scores than boys over the whole ability range. The greatest superiority is shown in tests of the basic skills rather than in tests of reasoning and the smallest differences occur in tests of mathematical or spatial ability. (*c*) There are appreciable differences in the achievement of children from different areas. (*d*) Some evidence suggested a relationship between the member of family listed first by the child and his own ability. A correlation of –0·2 was found between number of siblings and aggregate test score. (*e*) Factor analysis of the intercorrelations of sub-tests showed a consistent factor pattern for the three sets of tests.

I.—INTRODUCTION.

THE need for a continuous process of educational guidance throughout school life has often been stressed (e.g., Wiseman 1964, p. 150). One of the necessities for an adequate guidance programme is the availability of appropriate tests of ability and attainment. As far as the primary school is concerned, there is no shortage of tests of various types and for various ages, but one of the problems facing the teacher is the basic one of comparability of scores. It is clear from the work done in psychometrics over the past decades that adequate test-standardisation is difficult to achieve, mainly because of significant regional differences in ability and attainment. The comparison of results from tests of different origin, standardised at different times and in different places, is a hazardous exercise. What the teacher requires is a set of tests which have all been standardized at the same time and on the same population. This was the major aim behind the construction of the France-Wiseman Educational Guidance Programme.

The first draft of the Programme was prepared in the Autumn of 1963, in the form of three booklets intended to be worked by children in the first, second and third year of the junior school course. To provide try-out and item analysis data a total of 500 children in each age-group worked through the try-out booklets without a formal time limit in schools in Caithness, Aberdeen, Londonderry, Anglesey, Bury, Wallasey and Hampshire.* The item analysis eliminated items which did not discriminate adequately, or showed a significant sex difference. The items were arranged in order of difficulty within each sub-test. The make-up of sub-tests in the final edition of the Programme is now identical for each of the work-books (1A, 2A and 3A):

Comprehension: Questions on a number of passages.

Vocabulary: Choice of one out of five definitions for each of fourteen words.

Verbal Reasoning: Choosing one out of five or six words to complete a group of five words.

Spatial Reasoning: Choosing one out of five or six diagrams to complete a group of five diagrams.

Addition/Subtraction: Fifteen examples of each.

Likes and Dislikes: Children are asked to list the things they like and dislike and to say what kind of work they would like to do when they grow up.

Picture Composition: Children are asked to draw any picture in any medium they choose on the centre double page.

Members of Family: Children are asked to write the names of members of their family within a square.

Multiplication/Division: Fifteen examples of each.

Spelling: Twenty-six sentences each with a missing word, to be completed as the teacher reads the sentence aloud.

Reading: Thirty-six words to be read to the teacher.

Mechanical Arithmetic: (+ –x ÷) Thirty examples.

Arithmetic Problems: (A) Includes weights and measures. (B) Money. (C) Number Series.

Composition: Children are asked to complete a simple story.

The back cover of each work-book incorporates a flap to which marks can be transferred page by page to facilitate additions of raw scores. It is intended that ten of these totals should be translated to normalised standard scores (M=100, s.d.=15) and aggregated in four ways:

Verbal Skills: (i) Spelling; (ii) Reading.

Verbal Reasoning and Comprehension: (iii) Composition; (iv) Comprehension (including Vocabulary); and (iv) Verbal Reasoning.

Mathematical Skills: (vi) Number.

Non-Verbal Reasoning: (vii) Numerical Reasoning; (viii) Spatial Reasoning; (ix) Pictorial Composition.

And Grand Total: (x).

The standard scores for each child derived from tables in the teacher's book, are then transferred to the specially designed record card on which the items are arranged in the same order as on the flap.

---

* We are grateful to the authorities concerned, and to the teachers in the try-out schools, for their most helpful co-operation.

For the standardisation, the sample of one in four primary schools was selected at random by choosing the fourth, eighth, twelfth, etc., school from an alphabetically arranged list of primary schools in Solihull and Warwickshire. The head teachers of these schools attended a briefing conference, and the teacher's book of instructions for administering and marking the work-books was made available to all participating schools well in advance.

There were rather more than 2,000 children tested in each of the age groups. The testing took place in November, 1964, after the two older groups had each worked the Moray House Verbal Reasoning Test which is a normal part of the current Warwickshire procedure for transfer from primary to secondary school.

Reliability of the Programme.

The teachers were asked to record separately the number of incorrect odd-numbered and even-numbered items. The whole tested population was thus used to calculate the reliability coefficients by using the ' split-half ' method and the Spearman-Brown prophecy formula :

Work-book	Age-group	Number of Children	Reliability Cofficient
1A	1955–6	2,354	0·980
2A	1954–5	2·199	0·966
3A	1953–4	2,114	0·978

II.—STANDARDISATION OF TEST SCORES AND THE EFFECT OF MONTH OF BIRTH.

For this purpose, each age-group was divided by month of birth into the three termly groups with months of birth September to December, January to April, and May to August, since these are more closely related to the terms of the school year than are the four quarters of the year.

A frequency distribution table was prepared for each of the eleven major variables, for each of the three age-groups and for each of the three termly groups, as defined above, within each age-group. Cumulative frequencies were calculated and these were converted into percentile levels, which were translated into normalised standard scores with a mean of 100 and an S.D. of 15. Raw scores were now plotted against these equivalent normalised standard scores. Finally, the raw scores were read off for each termly group at intervals of five points of standard score. Then by interpolation and extrapolation, the tables were prepared for the conversion of raw scores to normalised standard scores at two-monthly intervals. It is intended that the standard scores should be restricted to multiples of five.

A useful by-product of this graphical and geometrical method is the visual check presented firstly as to whether or not any set of raw scores is normally distributed in the population and, secondly, as to the size and nature of the differences of the distribution of scores between the three termly groups and between boys and girls.

All the tests were worked without any time limit. The most notable difference between timed or untimed tests is that untimed tests produce (*a*) rather higher scores than might have been expected in a normal distribution at the lowest end of the ability range (i.e., standard score below the 75 level) and (*b*) rather lower scores at the highest level (i.e., above 125). This is understandable in the case of an untimed test since in the latter case there must needs be a compromise between the need to keep the length of the test within reasonable bounds and the need to stretch the most able children. In most group test situations it is not usually necessary to provide a ' vernier-like ' differentiation between individuals in the higher 5 per cent. of the child population.

Another impression derived from a study of the graphs is that the difference in achievement between Autumn-born (September to December) children and Spring-born (January to April) children is rather less than that between Spring-born children and Summer-born (May to August) children. Summer-born children have little chance of being admitted to infant schools in general, until after their fifth birthday. This means that by November, 1964, for instance, many of the children born in the Summer of 1956 would only have completed nine terms in school. On the other hand, many children born in the Spring of 1956 could have had eleven terms in school and those born in the Autumn of 1955 might have had four complete school years. It was clear, too, that the differences of achievement for term of birth were greater for the 8+ group than for the older groups.

A study of the distribution lines showed, too, that the effect of term of birth was not uniform over the whole range of ability. There was a suggestion that the intermediate Spring-born group had similar achievements to the older Autumn-born children at the top of the ability range but at the bottom of the ability range there was a greater similarity between the Spring-born and the Summer-born. It is possible that these affinities reflect the basis of streaming or grouping within the year-groups contained.

Table 1 records the differences of achievement between the three termly groups for the major variables. Separate differences are here (as elsewhere in this paper) expressed to the nearest whole number: an attempt to record differences to a greater degree of accuracy is not justified.

TABLE 1

DIFFERENCES OF ACHIEVEMENT IN TERMS OF STANDARD SCORE BETWEEN (*a*) AUTUMN-BORN AND SPRING-BORN, AND (*b*) SPRING-BORN AND SUMMER-BORN CHILDREN.

Major Variable	1A (8+ age-group)		2A (9+ age-group)		3A (10+ age-group)		Average of 1A, 2A and 3A		Av. Term difference
	(*a*) Au-Sp	(*b*) Sp-Su	(*a*) Au-Sp	(*b*) Sp-Su	(*a*) Au-Sp	(*b*) Sp-Su	(*a*) Au-Sp	(*b*) Sp-Su	
(i) Spelling	2	4	3	1	1	2	2·0	2·3	2·2
(ii) Reading	2	3	2	1	1	1	1·7	1·7	1·7
(iii) Number	4	3	2	2	1	3	2·3	2·7	2·5
(iv) Composition	2	4	3	2	2	1	2·3	2·3	2·3
(v) Comprehension (inc. Vocabulary)	4	4	3	1	2	3	3·0	2·7	2·9
(vi) Verbal Reasoning ..	2	4	2	2	0	0	1·3	2·0	1·7
(vii) Numerical Reasoning	4	4	2	2	1	2	2·3	2·7	2·5
(viii) Spatial Reasoning ..	1	4	1	2	1	3	1·0	2·3	1·7
(ix) Pictorial composition	1	3	2	1	2	1	1·7	1·7	1·7
(x) Grand Total	3	5	3	1	1	3	2·3	3·0	2·7
Average for (i) to (ix)	2·4	3·7	2·2	1·6	1·2	1·8	2·0	2·3	2·2
Average term-difference ..	3·1		1·9		1·5		2·2		

(*a*) Autumn-born = September to December.
(*b*) Spring-born = January to April.
(*c*) Summer-born = May to August.

From Table 1 we can clearly see (*a*) the tendency for the age difference to decrease as children grow older (3·1, 1·9 and 1·5); (*b*) that differences are generally greater for Spring and Summer-born than for Autumn and Spring-born children (2·3 and 2·0); and (*c*) that the differences due to age are greatest for Comprehension (2·9), Numerical Reasoning (2·5) and Number (2·5) and least for Reading (1·7), Verbal Reasoning (1·7), Spatial Reasoning (1·7) and Pictorial Composition (1·7). These differences are hardly large enough to allow firm conclusions to be drawn. Experience elsewhere, however, suggests that one can expect smaller age allowances with spatial tests and other activities which are not so closely allied to the everyday work of the classroom. Table 1 gives some indication of a link between the amount of age allowance for a test and the closeness with which it matches normal work in school.

III.—SEX DIFFERENCES IN ABILITY AND ATTAINMENT.

Discussions on differences of achievement between boys and girls are most usually concerned with differences in mean performance for the whole age-group. The differences of spread in the achievement of boys and girls tends to be overlooked. It is the combination of average and spread which really needs to be considered. Plotting raw scores against normalised standard scores produces a straight line graph if the distribution of raw scores is normal The usual picture which results from plotting the raw test scores of boys and girls in the same graph is a pair of straight lines with the line for boys inclined at a steeper gradient than that for girls. The gradient is directly proportional to the standard deviation of the distribution. This means that the spread of raw scores for boys is almost always greater than the spread for girls. A direct consequence of this is that the superiority of the achievement of girls will be more marked at the lower levels of ability and less marked or even reversed, at the upper levels. It is not unreasonable, therefore, to expect more boys than girls to occupy places in special schools for educationally subnormal pupils and to find, at the other end of the ability scale, a larger number of boys than girls capable of following a course of university standard. Social and environmental factors are clearly important in determining and emphasising these differences of achievement between boys and girls.

Table 2 below shows that, over the whole range of ability, girls achieve higher scores than boys by something of the order of two to three points of standard score. The table also shows that not only is there a variation of this sex difference with level of ability, but also with type of activity. Girls tend to achieve higher scores relative to boys in tests which are more verbal and less spatial and which involve to a greater degree the application of skills learned rather than the capacity to reason. Even the two subjective assessments made by teachers (Composition and Pictorial Composition) follow this same kind of trend. Similarly, it is interesting to notice that the two Moray House Verbal Reasoning Tests show slightly greater sex differences than either the 2A or 3A work-books in the Programme.

From Table 2 it can clearly be seen that the greatest superiority of the achievement of girls is shown in Composition (5·4), Spelling (4·2) and Number (3·3). Boys manage to reduce the lead of the girls particularly in Pictorial Composition (—0·4), Numerical Reasoning (—0·1) and Spatial Reasoning (0·8).

These differences are related to the content of the tests in much the same way as those found in an earlier analysis of results on a different test battery in Kent (France 1963, pp. 235-240). A random 1 in 15 sample numbering about 1,200 of the Kent 1945 age-group, with a median age of 10 : 7, was tested in

TABLE 2

STANDARD SCORES ACHIEVED BY GIRLS IN EXCESS OF THOSE ACHIEVED BY BOYS AT THREE LEVELS: (*a*) At one standard deviation below the mean (85).
(*b*) At the mean (100).
(*c*) At one standard deviation above the mean (115).

Major Variables	1A (8+ age-group)			2A (9+ age-group)			3A (10+ age-group)			Average of 1A, 2A and 3A			Average all levels
	(*a*)	(*b*)	(*c*)	(*a*)	(*b*)	(*c*)	(*a*)	(*b*)	(*c*)	(*a*)	(*b*)	(*c*)	
(i) Spelling	5	4	2	6	5	2	5	5	4	5·3	4·7	2·7	4·2
(ii) Reading	5	3	1	4	2	1	3	2	1	4·0	2·3	1·0	2·4
(iii) Number	2	2	2	4	4	4	4	4	4	3·3	3·3	3·3	3·3
(iv) Composition	7	6	5	5	5	5	5	5	6	5·7	5·3	5·3	5·4
(v) Compehension (inc. Vocabulary)	4	2	0	4	2	0	1	–1	–2	3·0	1·0	–0·7	1·1
(vi) Verbal Reasoning	5	3	1	4	2	2	1	0	–1	3·3	1·7	0·7	1·9
(vii) Numerical Reasoning	1	0	–2	1	0	–2	1	0	0	1·0	0	–1·3	–0·1
(viii) Spatial Reasoning	–1	–1	–1	2	2	1	1	2	2	0·7	1·0	0·7	0·8
(ix) Pictorial Composition	1	0	–1	1	0	0	–1	–2	–2	0·3	–0·7	–1·0	–0·4
(x) Grand Total	3	2	0	5	4	1	3	3	3	3·7	3·0	1·3	2·7
Average for (i) to (ix)	3·2	2·1	0·8	3·4	2·4	1·4	2·1	1·7	1·3	2·9	2·1	1·2	2·1
Average all levels	2·0			2·4			1·7			2·1			
Moray House Verbal Reasoning	—	—	—	4	3 3·0	2	3	2 2·7	3	3·5	2·5 2·8	2·5	2·8

January, 1956. The results are shown below in Table 3 alongside the Warwickshire 10+ age-group results in 3A for comparison.

TABLE 3

STANDARD SCORES ACHIEVED BY KENT GIRLS (AGED 10+) IN EXCESS OF THOSE ACHIEVED BY BOYS AT THREE LEVELS: M—s.d.=85, M=100; M+s.d.=115.

Type of Test	Kent			Averages	
	85	100	115	Kent (10+)	Warwick (10+)
(i) Spelling	7	3	1	3·7	4·7
(ii) Reading	6	1	0	2·3	2·0
(iii) Number	4	2	1	2·3	4·0
(v) Comprehension	7	1	—1	2·3	—0·7
(vi) Verbal Reasoning	1	0	0	0·3	0·0
(vii) Numerical Reasoning	5	0	—2	1·0	0·3
(viii) Spatial Reasoning	—3	—4	—7	—4·7	1·7
Average for all tests	3·9	0·4	—1·3	1·0	1·7

Thus, three kinds of tests, Spelling, Reading and Number tend to favour girls more than boys in both Warwickshire and Kent. Similarly, the Verbal and Numerical Reasoning tests do not produce much difference in scores for boys and girls. The results from the two counties, however, do not support each other for the Comprehension and Spatial Reasoning tests. This last fact is not really surprising in the light of the known differences of type of content between the two test batteries. A diagram of these sex differences showing the Kent and Warwickshire results plotted along two axes at right angles suggests a greater agreement than might at first sight appear from Table 3. In fact, a very good line of fit for all the tests except Comprehension and Spatial Reasoning can be defined by:

$$W=\frac{5K}{3}-1$$

where W and K represent the superiority in points of standard score for girls in Warwickshire and Kent on their respective test batteries.

Table 3 suggests that a dichotomy of the type of test, reasoning *versus* skills, may have as much significance for the amount of the sex difference as the dichotomy verbal *versus* mathematical. If we make this four-fold classification we obtain the results summarised in Table 4.

TABLE 4

AVERAGE STANDARD SCORES ACHIEVED BY GIRLS IN EXCESS OF BOYS FOR WARWICKSHIRE AND FOR KENT.

	Tests of Skills		Tests of Reasoning	
	Warwick	Kent	Warwick	Kent
Verbal	3·3	3·0	—0·3	1·3
Mathematical ..	4·0	2·3	1·0	—1·8

This lends particular weight to the supposition that the greatest superiority of girls is to be found in the exercise of skills rather than in the application of reasoning, and in tests with a verbal rather than with a mathematical content. Wisenthal (1965) has pointed out that girls show a higher achievement than boys in both the United States and Canada as well as in the United Kingdom. He goes on to emphasise the link between measure of attitude to school work and achievement as shown by test scores.

IV.—REGIONAL DIFFERENCES.

The distributions of Grand Total standard scores for each of the workbooks 1A, 2A and 3A were plotted separately for each of the five areas of Warwickshire and for the County Borough of Solihull, and comparisons were made with the level of standard scores for the whole country. An independent check was available through the Moray House Verbal Reasoning Test, which had been taken by some of the Warwickshire third-year children, but by nearly all children in the fourth junior year. Table 5 shows the details area by area.

TABLE 5

STANDARD SCORES ACHIEVED BY CHILDREN IN DIFFERENT AREAS OF WARWICKSHIRE IN EXCESS OF THE COUNTY AVERAGE.

Area	Approx. number in age-group	1A 8+ age-group	2A 9+ age-group	3A 10+ age-group	Average of 1A, 2A, 3A (F–W)	Moray House V.R. 10+ age-	Average of F–W and MHVR
Solihull	330	+5	+7	+8	+6	+6	+6
Sutton Coldfield	220	+6	+4	+7	+6	+7	
Rugby	270	+3	+6	+6	+5	+5	
Nuneaton	250	–4	–2	–4	–3	–2	–3
Mid-Warwickshire	380	–1	0	–3	–2	–4	
Central Area	800	–2	–5	–3	–3	–3	
Total (Whole County)	2250	0	0	0	0	0	0

There is a marked difference in levels of ability between children at school in Solihull, Sutton Coldfield and Rugby, on the one hand, and those at school in the rest of Warwickshire. As Table 5 shows, the order of difference between the two main regions is almost ten points of standard score. A knowledge of the magnitude of such differences is naturally important to the local education authority in planning its provision of educational facilities if it is to provide equitable treatment for different parts of the county.

Similar differences between regions were found in Kent (France 1963, pp. 318-320), but in this case the transition from areas of high talent to areas of low talent was much less abrupt. In fact, it has been possible to produce a map of Kent using the analogy of the weather map in which 'iso-ints' (lines joining places with an equal proportion of children scoring more than 110 on the intelligence test) take the place of isobars.

Table 6 shows the variation of the differences of standard score between the areas for each of the major variables. On the whole, it is the assessments for Composition and Pictorial Composition which show the smallest differences between the Solihull, Sutton Coldfield, Rugby area and the Nuneaton, Mid-Warwickshire and Central area. The largest differences for the three age-groups are shown in the standard scores for Numerical Reasoning and Number, but differences for the remaining sub-tests are almost as large. The differences in attainment between the central and suburban areas of the Manchester conurbation demonstrated by Wiseman (1964) are thus shown to have their counterpart when county areas are examined. The implications of such large regional differences are not always fully realised by educational planners and administrators, nor do we yet know enough about their origins to be capable of adjusting our educational techniques accordingly. But such facts alone provide powerful support for the necessity for systematic educational guidance in all schools and all regions.

TABLE 6

MEDIAN STANDARD SCORES ACHIEVED BY CHILDREN IN DIFFERENT AREAS IN EXCESS OF THE COUNTY MEDIAN STANDARD SCORE.

(Average of 1A, 2A and 3A, 8+, 9+ and 10+ age groups)

	AREA						
Major Variable	Solihull (5)	Sutton Cold-field (4)	Rugby (3)	Nun-eaton (2)	Mid-War-wick (1)	Central Area (6) to (9)	Average Differ-ence (5+4+3) — (2+1+6/9)
(i) Spelling	+4	+4	+3	−3	−3	−3	7
(ii) Reading	+3	+3	+2	−1	−2	−2	5
(iii) Number	+5	+5	+2	−3	−2	−6	8
(iv) Composition	+2	0	0	−2	−2	0	2
(v) Comprehension (inc. Vocabulary)	+4	+5	+2	−6	−2	−3	8
(vi) Verbal Reasoning ..	+4	+4	+3	−4	−2	−3	7
(vii) Numerical Reasoning	+3	+6	+2	−6	−3	−5	9
(viii) Spatial Reasoning .	+3	+5	+3	−4	−4	−3	8
(ix) Pictorial Composition	+1	+1	+1	−2	−3	+2	3
Average of (i) to (ix) ..	+3·2	+3·7	+2·0	−3·4	−2·6	−2·6	5·9
Global Average	+3·0			−2·9			5·9

V.—THE NATURE OF FAMILY RELATIONSHIPS.

The results reported in this section are based on an analysis of the information derived from page 18 in each work-book. On this page a blank square (4-in. × 4-in.) had this instruction above it: " Write the names of the members of your family inside this square."* Teachers were asked to give no additional explanation and, in reply to any question, to say " Just do what is asked in the way you think best."

The analysis of the member of family placed *last* showed relatively small differences for the various family relatives. About 30 per cent. of children placed themselves last. The differences between relatives were, however, very much magnified for the member placed *first*, but there was close agreement between the trends for members placed first and last. Table 7 below shows the percentages of children placing different members of family first and last. All figures are rounded off to the nearest whole number to facilitate comparison.

One interesting feature of Table 7 is the trend of change from the 8+ age group to the 10+ age group. The percentage of children placing father or mother first increases as the child grows older. Correspondingly, there is a decrease in the percentage of children placing a brother or sister or themselves first in their list. These changes are confirmed by the evidence of that part of Table 7, showing the member of family placed last.

---

* The idea of asking children to list names in this way arose out of a paper read by N. J. Bradley, University of Leicester Institute of Education, at the Education Section meeting of the British Psychological Society, in 1963.

TABLE 7

PERCENTAGES OF BOYS OR OF GIRLS AND THE PLACEMENT THEY ACCORD TO MEMBERS OF FAMILY.

Member of Family Placed First by	Boys / Girls	Placed First			Placed Last		
		1A (8+ age-group)	2A (9+ age-group)	3A (10+age-group)	1A (8+ age-group)	2A (9+ age-group)	3A (10+age-group)
1. Father	Boys	29	33	35	8	7	5
	Girls	20	24	26	10	8	9
2. Mother	Boys	32	33	36	8	7	5
	Girls	46	46	50	5	3	3
3. Brother	Boys	12	10	9	11	12	13
	Girls	9	8	6	8	11	11
4. Sister..........	Boys	8	8	6	11	10	12
	Girls	11	10	7	11	8	9
5. Other male	Boys	2	2	1	10	12	13
relative	Girls	1	2	1	11	17	12
6. Other female	Boys	1	1	2	8	11	9
relative	Girls	2	2	1	10	13	13
7. Self	Boys	14	12	9	30	28	31
	Girls	10	7	7	33	30	31
8. None of these	Boys	2	1	2	14	13	12
	Girls	1	1	2	12	10	12
Totals..........	Boys	100	100	100	100	100	100
	Girls	100	100	100	100	100	100

As might have been expected, boys and girls behave very differently in this situation. The proportion of boys placing their father first is about the same as the proportion of boys placing their mother first. On the other hand, the proportion of girls placing their mother first is almost double the proportion of girls who place their father first. Those children who place a brother or sister first tend to choose a sibling of the same sex.

A scrutiny of the analysis of the data suggested some difference in ability and attainment between children who made different choices for first place in their list of members of their family. Attention was concentrated on the three groups of children who gave first place (i) to their father, (ii) to their mother, and (iii) to themselves. The criterion adopted was the Grand Total of all raw scores on the Programme and the result is summarised in Table 8.

Thus, the three groups of boys placing their father, mother or self first, respectively, can be placed in that order of overall ability. For the girls, the two groups placing father and mother first are of about the same level of ability, but the group of girls listing themselves first seem to consist of girls of a lower level of ability.

For the two groups of (8+) boys choosing father and mother for first place, a further analysis was made to discover any particular content bias. This ' father-first ' group showed a superiority of 2 to 3 points of standard score over

TABLE 8

ATTAINMENT OF CHILDREN RELATED TO MEMBER OF FAMILY PLACED FIRST.

Excess of recorded standard score over the true standard score level within each group. For groups of children placing first (i) Father (ii) Mother, and (iii) Self; at three levels in each group.

		1A (8+ age-group)			2A (9+ age-group)			3A 10+ age-(group			Average for 1A, 2A and 3A			Average all levels
		85	100	115	85	100	115	85	100	115	85	100	115	
(i) Group placing	Boys	3	2	3	0	1	1	3	3	2	2	2	2	2·0
FATHER first......	Girls	3	4	3	5	3	2	4	5	2	4	4	2	3·4
(ii) Group placing	Boys	–1	1	1	–2	–1	–1	–1	–1	–1	–1	0	0	0·7
MOTHER first	Girls	5	4	2	5	4	1	4	3	4	5	4	2	3·6
(iii) Group placing SELF	Boys	–4	–5	–1	–4	–2	0	–5	–1	0	–4	–3	0	–2·4
first	Girls	–6	–4	–2	–2	–4	–2	2	2	–1	–2	–2	–2	–1·9

the 'mother-first' group of boys for Spelling, Reading, Comprehension, Numerical Reasoning and Spatial Reasoning. On the other hand, the 'mother-first' group of boys showed a superiority of 2 to 3 points of standard score for Composition, Verbal Reasoning and Pictorial Composition. There is a hint, therefore, that a measure of creative ability might prove to be linked more closely with a boy-mother than with a boy-father relationship.

VI.—SIZE OF FAMILY.

In this section the variable considered is strictly the number of brothers and sisters each child in the sample records on page 18 of the work-book.

Distribution of Family Size.

The investigation showed that the distribution of the logarithm of family size was almost exactly a normal distribution for each of the three age-groups. This was found to lead to the relationship:

$$S = \frac{400}{7} \log_{10}N+80$$

where N = the number of brothers and sisters recorded by each child
and S = the 'standard score' equivalent to the corresponding percentile levels for less than this number of siblings.

Thus, to find the percentage of children with less than two siblings, we substitute N = 2 and obtain

$$S_2 = \frac{400}{7} \log_{10}2+80$$

$$S_2 = \frac{400}{7} (0{\cdot}3010) + 80 = 17 + 80 = 97$$

Now 97 corresponds to a percentile level of 42 per cent. This means that 42 per cent. of the age-group have less than two siblings. Table 9 shows the values calculated by the formula.

TABLE 9

DISTRIBUTION OF SIZE OF FAMILY
Expressed in Standard Scores and equivalent Percentile Levels)

	NUMBERS OF BROTHERS AND SISTERS								
Less than	1	2	3	4	5	6	7	8	9 or more
	80	97	107	114	120	124	128	132	135
(Corresponding % level) :	9%	42%	68%	82%	91%	95%	97%	98%	99%

TABLE 10

COEFFICIENTS OF CORRELATION BETWEEN MAJOR VARIABLES AND NUMBER OF BROTHERS AND SISTERS.

(N=1,000±150. Standard Error for r=0·1/0·2 is of the order s.e.=0·03).
(All coefficients are negative)

Major Variable	1A (8+ age-group)		2A (9+ age-group)		3A (10+ age-group)		All ages 1A, 2A and 3A		Average for B. and G.
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	
(i) Spelling	·11	·19	·18	·21	·16	·18	·15	·19	·17
(ii) Reading	·14	·18	·23	·21	·13	·14	·17	·18	·18
(iii) Number	·14	·14	·20	·24	·11	·22	·15	·20	·18
(iv) Composition	·09	·09	·17	·18	·16	·14	·14	·14	·14
(v) Comprehension (inc. Vocabulary)	·16	·21	·23	·25	·16	·20	·18	·22	·20
(vi) Verbal Reasoning ...	·13	·17	·16	·21	·17	·20	·15	·20	·18
(vii) Numerical Reasoning	·14	·17	·21	·23	·15	·18	·17	·19	·18
(viii) Spatial Reasoning ..	·11	·16	·13	·23	·14	·12	·13	·17	·15
(ix) Pictorial Composition	·07	·04	·08	·06	·06	·06	·07	·05	·06
(x) Grand Total	·16	·17	·24	·26	·15	·21	·18	·21	·20
Moray House V.R. I.Q.	—	—	·24	·22	·11	·16	·18	·19	·19
Average of (i) to (ix) above	·12	·15	·18	·20	·14	·16	·15	·17	·16
Boys and Girls.	·14		·19		·15		·16		

The standard scores so derived differ from the observed values for children in each of the three age-groups usually by not more than one point of standard score. It will be noticed from the table, for instance, that about 9 per cent. of children are without brothers and sisters.

Family Size and Attainment.

For each of the ten or eleven major variables and for each of the three age-groups, 8+ (1A), 9+ (2A) and 10+ (3A), the coefficient of correlation was calculated for boys and girls separately, with the size of family, defined as the number of brothers and sisters possessed by each child in the sample. It was seen above that the logarithm of the number of brothers and sisters is distributed normally so that care needs to be used in interpreting the absolute magnitude of the derived coefficient of correlation.

The chief value of the results contained in Table 10 below lies in the comparison which can be made between the three age-groups, between boys and girls and between the different major variables.

All correlations, as might be expected, were found to be negative. Signs have been omitted so that the correlation coefficients in Table 10, in fact, relate achievement with smallness of family.

Table 10 show sthat family size and attainment are significantly related with an overall correlation of the order of –0·2. This agrees closely with the results of other investigations of the relationship between I.Q. and family size, where correlations range between –0·2 and –0·3 (see Nisbet, 1953). The particular sub-test which is most closely related to family size is that for Comprehension (including Vocabulary) with an average correlation of –0·20. This is of about the same order of agreement as that shown by the Moray House Verbal Reasoning Test (–0·19). It is particularly interesting to note that the sub-tests which are least affected by family size are Pictorial Composition (–0·06), and Composition (–0·14) and that these were the tests found to be least affected by regional differences (Table 6).

VII.—INTER-CORRELATIONS OF SUB-TESTS AND FACTOR PATTERN.

In order to reduce the magnitude of this operation, boys and girls were considered together. Previous experience suggests that similar trends in the correlations are seen for boys and girls. Generally, correlations for boys are found to be slightly higher than the corresponding correlations for girls.

Table 11 shows that the separate sub-tests which correlate most highly with the Grand Total for all age-groups are Comprehension (including Vocabulary) (0·89) and Numerical Reasoning (0·87). On the other hand, the sub-tests which correlate least with the Grand Total are Spatial Reasoning (0·64) and Pictorial Composition (0·25). It can be seen that the Pictorial Composition correlations are appreciably lower than those of any other of the major variables. In fact, the highest correlation shown is 0·33 and this is with the only other subjective assessment, i.e., Composition.

TABLE 11

COEFFICIENTS OF CORRELATION BETWEEN ALL PAIRS OF THE MAJOR VARIABLES.

1A=8+ age-group : N_1=2,300
2A=9+ age-group : N_2=2,200
3A=10+ age-group : N_3=2,100

(Decimal points have been omitted)

		(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)	MH	Av.
(i) Spelling	1A		85	64	75	70	65	66	42	23	84	—	613
	2A		85	66	76	57	67	69	46	18	83	69	605
	3A		83	66	74	68	69	72	53	19	84	76	630
(ii) Reading	1A			65	81	74	60	71	42	24	85	—	628
	2A			66	79	58	69	69	48	18	86	69	615
	3A			58	78	65	64	68	52	19	84	72	609
(iii) Composition	1A				64	58	50	59	40	36	66	—	545
	2A				69	53	57	62	47	32	69	63	565
	3A				61	57	56	58	46	32	66	55	543
(iv) Comprehension and Vocabulary	1A					76	68	79	51	27	89	—	651
	2A					68	75	79	58	22	89	72	658
	3A					74	71	79	63	25	89	77	656
(v) Verbal Reasoning	1A						63	79	53	27	82	—	625
	2A						60	66	56	20	72	60	548
	3A						67	73	65	20	81	73	611
(vi) Number	1A							74	46	20	85	—	558
	2A							79	55	21	89	66	604
	3A							78	61	20	87	70	608
(vii) Numerical Reasoning .	1A								53	25	87	—	633
	2A								61	22	87	74	634
	3A								71	23	88	79	653
(viii) Spatial Reasoning	1A									23	57	—	438
	2A									19	64	58	488
	3A									23	71	62	543
(ix) Pictorial	1A										27	—	256
	2A										23	21	215
	3A										25	20	226
(x) Grand Total	1A											—	736
	2A											78	736
	3A											84	750
Moray House V.R. I.Q.	1A												—
	2A												613
	3A												639
Average of (i) to (ix) for 1A, 2A, 3A		616	617	551	655	595	590	640	490	232	741	626	554

Principal components analyses for the three separate age-groups were carried out on the Manchester University Atlas computer, followed by a varimax rotation. The loadings on eight factors are given in Table 12. It will be seen that the factor patterns for the three age-groups are closely similar, and that—

TABLE 12

FACTOR ANALYSIS (Varimax)

Factor-loadings for separate analyses of 1A (8+), 2A (9+) and 3A (10+) (decimal points omitted)

	I	II	III	IV	V	VI	VII	VIII
	1A 2A 3A	1A 2A 3A	1A 2A 3A	1A 2A 3A	1A 2A 3A	1A 2A 3A	1A 2A 3A	1A 2A 3A
(i) Spelling	86 85 77	26 23 32	14 00 02	20 19 22	30 20 26	13 20 21	15 14 16	08 06 06
(ii) Reading	70 81 84	27 22 18	27 22 28	27 19 17	22 27 21	20 04 06	14 16 19	09 06 07
(iii) Composition	30 36 30	87 83 87	14 11 12	16 17 16	19 18 19	13 11 09	15 17 15	18 17 17
(iv) Comprehension and Vocabulary	43 50 47	16 27 24	67 56 61	17 29 27	32 32 28	24 22 20	22 23 26	11 08 11
(v) Verbal Reasoning ..	36 27 33	21 16 22	20 10 16	76 86 78	28 21 25	24 13 14	24 23 29	11 08 07
(vi) Number	28 37 32	16 17 21	14 10 13	17 21 20	88 82 82	17 16 15	19 21 25	07 08 19
(vii) Numerical Reasoning	30 36 35	22 20 21	23 16 22	32 27 25	41 41 38	68 63 61	24 27 36	09 09 09
(viii) Spatial Reasoning ..	15 19 21	13 14 15	11 08 12	15 21 21	18 19 22	12 12 13	94 90 88	10 08 10
(ix) Pictorial Composition	07 06 06	13 10 11	05 03 04	06 06 04	06 06 05	05 03 03	08 06 07	98 99 98
Average % Variance ...	22·1	11·9	6·3	11·0	13·6	6·6	12·9	11·8

apart from Factor I—each factor is clearly identified with a particular sub-test. Factor I shows high loadings for both Reading and Spelling, and it should be noted that the Reading loadings increase with age, whereas those for Spelling decrease. The only other heavy loadings in the table which show a clear age-trend are those for Numerical Reasoning (Factor VI) and Spatial Reasoning (Factor VII), both of them showing a decrease in importance with increasing age.

The last factor listed (VIII) is the only one which is a true specific, with no significant loadings on any variable except Pictorial Composition—which variable, *mutatis mutandis*, shows no significant loadings on any other factor. The pattern of loadings in the other factors show relationships in the expected directions in most cases.

Factor I is not only one connected strongly with the verbal skills of Reading and Spelling, it also shows signs of being a more general educational factor with sizeable loadings on all variables but Spatial Reasoning and Pictorial Composition. Factor II is clearly labeled ' Composition ' with heavy loadings of over 18 in this variable. But its remaining variance is spread widely over the remaining sub-tests. In view of the subjective nature of the assessment of composition (marked by the teachers using general impression) one might suspect that this spread indicates halo effect. And yet the other purely subjective assessment—Pictorial Composition—shows exactly the opposite effect.

Factor V is the next most important one, with clear loadings on Number, predictably heavy loadings on Numerical Reasoning, and strong connections with Spelling, Reading and Comprehension. The Verbal Reasoning Factor (IV) has significant loadings in the other reasoning test, and also with Spelling, Reading and Comprehension. Factor III (Comprehension and Vocabulary) has the expected loadings on the other verbal tests, but an unexpectedly high loading on Numerical Reasoning—a connection which is reflected in Factor VI.

All in all, the factor analysis is highly satisfactory, inasmuch as it reveals a stable factor pattern from 8+ through to 10+, and the matching of individual sub-tests with factors is just the kind of structure one would hope for in a battery intended for the purposes of educational guidance.

VIII.—CONCLUSION.

The France-Wiseman Educational Guidance Programme was devised to meet a need. The present ferment of change in the organisation of secondary school education, and concern about individual backwardness and environmental handicap, are tending to spotlight the need for a system of educational guidance firmly rooted in the primary school. Few local education authorities, except the very largest, can have available the expert knowledge and experience needed to devise, to validate and to standardise the objective tests which form the basis for guidance. The emphasis throughout this Programme has been on the child as a whole. The teacher needs an instrument which is fully comprehensive, covering intellect and personality and which, at the same time, is quite compact. By the provision of a wide variety of short untimed tests, to be done whenever convenient and as part of normal school work, it is hoped to avoid undesirable backwash effects, while the provision of a specially designed record card ensures adequate recording of guidance data. From the psychologist's point of view, he needs to be satisfied that the standard scores produced by a child are really comparable. Few, if any, norms in present day use can be regarded as national norms. Comparison between test scores is difficult unless it is known that the tests have been standardised on the same population at the same period of time. It is far more valuable to the teacher and to the educationist to have available really comparable norms based on a population of known background of the order of size of 1,500 to 2,000 than to collate a mixed bag of populations numbering 20,000 and upwards.

There are much bigger differences in the geographical distribution of abilities than most people realize. Normally, an authority is only interested in the variations within its own boundaries. It is invidious to make comparisons against a hypothetical ' national ' yardstick, but so long as we have decentralised control of educational administration, then each authority needs to ensure that these differences are recognised and catered for inside its own frontiers. It is hoped that this Programme, restricted in its use to local education authorities, will meet a real need.

IX.—REFERENCES.

BLANDFORD, J. S. (1957). *Standardised Tests in Junior Schools with Special Reference to the Effects of Streaming on the Constancy of Results.* M.A. Thesis, University of London Library.

FRANCE, NORMAN (1953). *The School Record Card and the Teacher.* M.Ed. Thesis, University of Nottingham Library.

FRANCE, NORMAN (1963). *The Relationship Between Primary and Secondary School Success.* Ph.D. Thesis, University of London Library.

FRANCE, NORMAN (1964). The use of group tests of ability and attainment. *Brit. J. Educ. Psychol.*, 34, 19-33.

FRANCE, NORMAN, and WISEMAN, STEPHEN (1965). *F-W Educational Guidance Programme.* Collins, Glasgow.

NISBET, J. D. (1953). Family environment: a direct effect of family size on intelligence. *Occasional Papers on Eugenics*, No. 8. London: Cassell.

WISEMAN, STEPHEN (1964). *Education and Environment.* Manchester Univ. Press.

WISENTHAL, MILES (1965). Six differences in attitudes and attainment in junior schools. *Brit. J. Educ. Psychol.*, 35, 79-85.

Manuscript received 2nd June, 1965.)

RESEARCH NOTE

THE APPLICATION OF FACTOR ANALYSIS TO PROBLEMS OF LEARNING

By G. E. R. BURROUGHS

(*University of Birmingham*)

In an earlier article published in this *Journal* (Burroughs and Morris, 1962), the authors attempted to separate out and to define the different factors at work in the learning of a simple musical theme. Essentially, the procedure involved giving a sample of children eight trials of a musical task (the criterion task, C) and a battery of other musical tests (the reference tests, R) and proceeding to factor analyse the complete matrix of inter-correlations (criterion plus reference). We may symbolise this matrix

$$R = \begin{matrix} R_C & R_{CR} \\ R_{RC} & R_R \end{matrix}$$

and the factor loadings

$$\begin{matrix} F_C \\ F_R \end{matrix}$$

The loadings running through the trials F_C were used to indicate the number of factors operating in this learning situation, with the pattern of these loadings showing how much each factor changed throughout the eight trials ; the loadings running through the reference tests F_R were used to identify these factors.

The paper published in this *Journal* attracted a certain amount of interest, especially in the U.S.A., and was reported upon with varying degrees of favour in certain review journals.

The present note is to say that the method of analysis indicated above was wrong. Since others have used the procedure, both before and after the appearance of the above paper, we write to bring the error to their attention and to suggest an alternative procedure.

In our paper we reported no more than eight factors. This identity between the number of factors and the number of trials was purely fortuitous. We did, in fact, take out more and had we continued this procedure until we had extracted the maximum number of factors, this number being at the most the number of trials plus the number of reference tests, doubtless each of the factors would have shown some loadings of the trials. Logically, this seems to be impossible. What is happening factorially in the trials is contained wholly and completely within the sub-matrix R_C above. The addition of the reference tests aids the identification of these factors in R_C but cannot change their number or the pattern of their loadings Experimentally, for example, the reference tests might have been given a week after the trials. At most, therefore, the number of factors to be identified cannot exceed the number of trials ; with a communality factor model it will be less than this number. Further, for any one experiment, the loadings running through the trials ought to be the same no matter what size and type of reference battery is used ; these trial loadings should be those secured from an analysis of R_C above.

The alternative procedure suggested is to analyse R_C alone and to carry out an appropriate rotation. Thereafter, a triangular analysis of the complete matrix R will yield a factor matrix with zeros in that part of it where they are required. A rotation can then be effected of this full analysis so that the loadings through the trials part of it are identical with those found from the R_C analysis. This will yield a factor matrix of the form

$$\begin{array}{lcc} \text{Trials} & F_C & O \\ \text{Reference tests} & F_{RC} & F_R \end{array}$$

F_C will be used to find what is happening factorially in the trials ; F_{RC} will permit identification of these factors ; F_R constitutes that part of the reference battery variance which is unique, i.e., which has nothing in common with the criterion task.

A fuller note on this procedure can be obtained from the author. It is thought to have some general applicability to correlation matrices of the form stated.

BURROUGHS, G. E. R., and MORRIS, J. N. (1962). Factors involved in learning a simple musical theme. *Brit. J. Educ. Psychol.*, XXXII, 18-28.

Sir Cyril Burt

BOOK REVIEWS

A TRIBUTE TO SIR CYRIL BURT

Banks, Charlotte, and Broadhurst, P. L. (Edit.) (1965). ΣΤΕΦΑΝΟΣ *Studies in Psychology Presented to Sir Cyril Burt.* London: University of London Press, ; and New York: Barnes and Noble, Inc., pp. 283, 42s.

This book contains a series of articles written in his honour by former students, colleagues and friends of Sir Cyril Burt. They, and the University of London Press, with which he has been long associated, presented it to him when he became an octogenarian.

It is a considerable privilege for the editor of the *British Journal of Educational Psychology* to be allowed to be associated with the presentation of this tribute to a great psychologist who has been closely associated with the *Journal* since it started over thirty-five years ago. C. W. Valentine, the founder of the *Journal*, in a brief biographical sketch which begins this book, and which was the last article he wrote before he died, says something of the debt which it owes to Sir Cyril. I had the honour of knowing Valentine well and of working with him for a number of years. During this time I learned from him more fully how great this debt really is and since I have myself been editor, this knowledge has been reinforced on many occasions. No one has been called on for help as often as Burt and when asked to write articles, organise symposia, review books, act as a referee or give advice upon policy matters he has never been known to decline; a count from the list of Sir Cyril's publications at the end of this book of studies shows that he has written more than fifty articles in this *Journal.*

From its nature a book like this cannot be reviewed in quite the same way as can others. If it has a central theme, it is meant to reflect something of Burt's own wide interests in psychology. This, even the sixteen contributors cannot, of course, do completely and many of the papers may perhaps be better regarded as contributions characteristic of their writers in that they represent distillates of one of their major lines of thought and research over many years. Such would be the case of Vernon's paper, *The Personality System*, in which he outlines with his customary lucidity and scientific caution, the ideas which go to make his book, *Personality Assessment*. Other papers which seem to fit into the same category are Pear's *Social Ascent in England Today*, Guilford's *Implications of Research into Creativity*, Gardner Murphy's *The Nature of Man*, Mace's *Causal Explanations in Psychology*, Koestler's *Motivation*, Hamilton's *Prediction in Practice*, Wilkie's *Neurology and Free Will*, and Cattell's long article on *Higher Order Factor Structures.* Readers who know of the writings of these authors will be able to predict their probable content, and they will not be disappointed with what they find. Not all the papers will be equally enjoyed by all readers of course, but they make good reading in their respective fields for they have the ring of sincerity of experts who are writing on matters in which they are personally very much involved.

The rest of the papers are research articles on a number of themes. Broadhurst and Eysenck contribute the results of an investigation which shows the presence of general emotionality in rats. Charlotte Banks has a study of some 300 boys in their late teens who had been sent to a detention centre. Fraser Roberts and Elizabeth Sedgley report a comparison between the intellectual development of a group of premature as compared with full-time babies, and Gertrude Keir describes an intensive psychological study of the children from Tristan da Cunha who were brought to this country after the volcanic eruption in 1961, which forced them to abandon their homes.

Again, no one can expect to find all the papers equally to their taste, but few will fail to find something of considerable interest. All admirers of Sir Cyril Burt will wish to own a copy of this book, all the more so when they know that profit from its sale is to go to a Cyril Burt Prize for Psychology at the University of London, to be awarded for outstanding research into individual differences. L. B. Birch.

BRITISH PSYCHOLOGICAL SOCIETY (1966). *Children in Hospitals for the Subnormal: a Survey of Admissions and Educational Facilities.* Available from the British Psychological Society, Tavistock House South, Tavistock Square, London, W.C.1, price 3s. 6d.

This is a report of a working party, set up in 1962, by the English Division of Professional Psychologists of the British Psychological Society, to consider certain aspects of the 1959 Mental Health Act in relation to mental subnormality. After conducting a survey of subnormal adults, the working party turned their attention to children in hospitals for the sub-normal. They carried out a survey of the characteristics of children admitted to hospitals for the sub-normal in a single year, and obtained information about the children's previous educational history as well as about the educational facilities provided in the hospitals.

The report, which summarises the results of the enquiry and makes a number of recommendations for improving facilities for severely subnormal children, is clear, concise and admirably presented. Obviously, the terms of reference and the resources of working parties of this kind are limited, but the survey, carried out by means of questionnaires and necessarily restricted to hospitals employing psychologists, can be regarded as a useful preliminary to a full-scale national enquiry, which would help to fill in the gaps in our knowledge which still exist. In particular, it would be valuable to have more detailed information than the report gives us on the incidence and nature of behaviour or emotional disturbance in children admitted to subnormality hospitals, and we need to know more about the high proportion of children (in this sample, 248 out of 403) who are not formally testable.

In pointing the way to further research, the report highlights some important practical issues. It provides sufficient evidence to indicate that for many of the children a hospital for the subnormal was not the most suitable placement. For example, of the total sample at least 24 per cent. had I.Qs. of 50, including 14 per cent. with I.Qs. over 70 and 4 per cent. with I.Qs. over 100. It was found, too, that although all the hospitals in the survey had school facilities of some kind, only a minority of the teaching staff had any formal qualifications. The links between local education authorities and hospital schools are shown to be tenuous, if they exist at all; the majority of educational psychologists, for instance, have little to do with severely subnormal children once they have been excluded from a local education authority school.

The members of the working party question the value of large subnormality hospitals and urge that small, family-type units should be provided for subnormal children who have to leave their families and are not suitable for admission to a residential school. They strongly recommend that the Department of Education and Science should be fully involved in the educational work of both junior training centres and schools in hospitals for the subnormal, and that better qualified teachers should be employed in hospital schools.

As recent work suggests that even severely subnormal children can benefit from appropriate educational stimulation, the working party have performed a valuable service in drawing attention to the failure to satisfy their educational needs. Few people concerned with subnormal children will not whole-heartedly support their recommendations. MAURICE CHAZAN.

BUROS, O. K. (Edit.) (1965). *The Sixth Mental Measurements Yearbook.* New Jersey: Gryphon Press, pp. 1714, $32.50.

For over a quarter of a century psychologists have relied heavily upon the *Mental Measurements Year Book* for information about the tests they wish to use. This new volume will immediately and inevitably find its place alongside the earlier ones. These do not, of course, become obsolete, but are supplemented by the latest volume and with it and the general index, *Tests in Print* (1961) by the same editor, form a unique and comprehensive body of knowledge about all known tests published in English.

The classification of tests and the general lay-out of this edition are essentially similar to the previous volumes. The Tests Section includes information about all tests published between 1959 and the middle of 1964. There are also a small number of older tests which were omitted from earlier Year Books, a number which appeared in earlier volumes but which have been reviewed again in the light of recent use and a number of older tests which are still attracting references in the testing literature. Altogether, there are some 1,200 tests included in the body of the text but the Index lists *all* tests known to be in print in 1964, over 2,000. These are cross referenced to *Tests in Print*, so that information about the older tests can easily be found in the earlier Year Books.

The salutory procedure of publishing reviews of the same work from more than one reviewer has been continued. Inevitably some reviews conflict and some are not very literate but on the whole they represent the best and most critical information available about currently obtainable tests. The bulk of reviewers are, naturally, American but rather more than 10 per cent. are British and a somewhat smaller number come from other English-speaking countries.

The Tests Section has continued to grow and in this edition it is about a third as big as the last. The Book Section, by contrast, is rather smaller in size though, in fact, over 8 per cent. more books are included. Size reduction has been achieved by publishing only a selection of the reviews which have appeared in the world's journals instead of including them all, as happened in some of the earlier Year Books.

This is not a book capable of an ordinary review. The series is so well known that to note its arrival is sufficient to ensure its inevitable purchase by every librarian dealing with the field of mental testing. One thing remains, to congratulate Professor Buros on the quality of the production and to thank him and his staff for continuing this tremendous service to psychometry. L. B. BIRCH.

Foss, B. M. (1963). *Determinants of Infant Behaviour, II.* London: Methuen and Co., Ltd., pp. xii+248, 45s.

The publication in book form of conference and seminar proceedings has become increasingly frequent in recent years. To be justifiable there should either be an attempt to present known facts or views in a new context or, alternatively, the book should contain reports of ongoing work or relatively new research findings. To be successful, the participants (or the editor) need to re-write their contributions since a spoken paper, which is interesting and easy to follow is rarely equally effective when read silently; conversely, it is difficult, and often dull, to listen to a paper prepared primarily for publication. Editing of the discussion is even more essential: to impose coherence, direction and shape at the time when it takes place requires not only an exceptionally skilled chairman, but also an unusually high level of co-operative and relevant interchange from all the speakers. To find both at one conference is exceedingly rare.

By these criteria few conference and seminar proceedings deserve to be permanently recorded in book form. The volume under review, though it has a number of valuable features, has not escaped some of the shortcomings from which such publications tend to suffer. Little attempt seems to have been made to link together the three parts of the book; the reader is given no guidance about which research techniques or which of the findings are new contributions to methodology and to knowledge; similarly he is left to tease out whether newly acquired insights invalidate or support previous results and what practical relevance, if any, the studies here discussed have for child-rearing or for clinical practice. In a brief introductory note, the editor states that " the discussions have been reduced to roughly one-tenth of the original volume "; but many points are left in mid-air; the contributions from the different disciplines and schools of thought are neither synthesised nor clearly differentiated; and the comments made by the discussants fail to illuminate the preceding paper or to convey the impression of a fruitful interchange.

On the positive side, many of the prepared contributions are interesting reports of current, and in some cases, long-term projects. Part I, which is the briefest, deals with animal studies. There are two papers, the major contribution being by Harlow on " The maternal affectional system " and its effects on the social and sexual development of young monkeys. Some of the results are by now widely known. A striking finding concerns the attitude of ' unmothered mothers ' to their own infants and the responses of the latter to maternal rejection. " One of the things that would unnerve experimenters was to watch the desperate efforts of these babies to make contact with the abnormal mothers. She would beat them and knock them down ; they would come back and make contact ; the mothers would rub their faces into the floor ; they would wriggle free and again attempt contact." Does this, perhaps shed some light on the affection not infrequently shown by children towards parents who ill-treat them as well as on what is being called the ' battered baby syndrome ' ?

Part II, entitled " Human Studies,' contains a regrettably brief account by Prechtl on " The mother-child interaction in babies with minimal brain damage." It is a pilot longitudinal follow-up of a small group of hyper-excitable children and a matched control group. If the preliminary finding are substantiated, they will help to decide whether certain behaviour problems should be explained in organic or psychogenic terms. Ainsworth reports " a short-term longitudinal study of a small sample of East African infants and their mothers," focussing on the development of infant-mother interaction. This section concludes with two papers by Wolff on the early development of smiling and on the natural history of a family. The first is based on the responses of eight babies during the first three months of life ; the second is a somewhat anecdotal account of one Boston Irish Catholic family and their five children.

Part III, " Method and Theory " is the least satisfactory because most fragmented section. There are five contributors, of whom three made short contributions (of less than six pages). The most interesting is a report by Shaffer. After summarising some preliminary results of a longitudinal study of the social behaviour of sixty infants during the first year of life, he considers some of the basic issues in the investigation of attachment behaviour. M. L. KELLMER PRINGLE.

HARPER, R. J. C., ANDERSON, C. C., CHRISTENSEN, C. M., and HUNKA, S. M. (Edit.) (1964). *The Cognitive Processes : Readings.* London : Prentice Hall, xii+717, 72s.

The task of a reviewer when faced with a book of readings is a curiously difficult one. Since almost all of the material has been published previously, as research papers, in learned journals, it would be inappropriate, and well nigh impossible, to offer a detailed critique of these. This reviewer has decided to examine the editor's aims, which are clearly stated in the preface, and to evaluate what they see as the Cognitive Processes, as exemplified in their selection of papers.

The editors of this volume are educational psychologists, professional teachers of teachers. In their view, contemporary psychological theory has not been developed to the point where it is of much value to the practitioner concerned with the manipulation of complex human behaviour. The most highly developed areas of psychology they see as being of little use in practice, a view shared no doubt by many in other applied fields.

The editors do not feel, predictably, that educational psychology should be abandoned until adequate theories are developed and, indeed, the book is an attempt to show how this dilemma might be tackled. The preface tells of an experiment, seemingly a successful one, in the teaching of psychology to Education students in the University of Alberta. Students in Alberta are required to attend a four semester course in Educational Psychology, the first two semesters of which are given over to General Psychology. The contents of this book provide the programme through which students worked in the last two semesters.

Reading the preface and being in some ways *set* to think in terms of *psychology for teachers* leads to rude awakening when the contents of the volume are examined. The first paper, which forms a section on its own and which the editors see as part of the preface, is Hebb's Presidential Address to the A.P.A., in 1960. This is a witty and scholarly examination of developments in the conceptualisation of problems which might broadly be categorised as to do with higher thought processes. The burden of Hebb's argument in this paper is that the serious analytical study of the thought processes cannot be put off any longer. What follows in the remainder of the book is a very thorough examination of the starting points, both empirical and theoretical.

The book is divided into six sections, containing some forty-seven papers by about as many authors. The editors provide one or two pages of introduction to each of these sections relating them often to Hebb's paper. The sections cover the following: Motivation, Neo-behaviouristic approaches to Cognition, Information processing, Computer Models, Cognition, Motivation and Personality and finally Cognitive Development. The area covered by these forty-seven papers is, of course, enormous: from physiological notions like Lindsley's 1957 paper on psychophysiology and motivation, through Hovland on the computer simulation of thinking, Festinger on cognitive dissonance to Roger Brown on the naming of objects by children.

One of the most valuable aspects of this book is its variety, not only of the problems raised, or the mixing of empirical and theoretical papers, but also in the various levels of explanation used. There are one or two papers of limited value which might have been profitably replaced but even these serve to broaden the conception of cognition.

The fact that this is intended to present a paradigm course for Education students, is surprising. In conception, a course such as this seems far wider and far bolder than that thought suitable for intending teachers. The editors have not, as yet, done a follow-up study on students passing through the course though from their experience they do feel able to assert its value. I doubt that many would question this claim.

Books of readings in different branches of psychology appear to be multiplying despite the fact that most of these seem to be of limited value and a sizeable proportion of no value. This book, however is a notable exception. The volume should be of considerable value, and not merely a value of convenience, to both Psychology and Education students. KEVIN CONNOLLY.

KAHN, R. L., WOLFE, D. M., QUINN, R. P., DIEDRICK SNOEK, J., and ROSENTHAL, R. A. (1964). *Organisational Stress: Studies in Role Conflict and Ambiguity*. London: John Wiley and Sons, pp. xii+470, 60s.

This book is another contribution to the long series of studies from the University of Michigan, concerning the behaviour of people in organisations. The intentions of the five authors were to determine the degree to which members of organisations are subject to conflicting or ambiguous role expectations; and to explore the dynamics of situations of high role conflict. How, for instance, are the role expectations communicated, and what coping responses does the role-holder utilise?

The authors acknowledge that they have been much influenced by Gross *et al's* classic *Studies in the School Superintendency Role*, but the present study is focussed on organisational roles in general, rather than any one specific role. The data are derived from two sources. First, an interview survey of 725 organisational employees was conducted, based on a national survey of adult Americans. Second, a more intensive interview study was undertaken focussed on fifty-three managers and supervisors in six industrial organisations.

The findings from the survey data are rather unimpressive. Apparently, 88 per cent. of the sample report some conflict between the ways different people require them to work. The major party to these conflicts tended to be the immediate superior or higher authority more generally. An equally high percentage of respondents perceived some ambiguity in the ways they were expected to behave by others.

High reported role conflict and ambiguity were associated with high tension and low job satisfaction. However, we can derive from these findings no idea of the intensity of the conflicts reported nor of whether they were even perceived as conflicts by those who worked with those interviewed.

The intensive study is much more worthy of note, since the authors utilise in it an objective measure of role conflict. Each interviewee (' focal person ') described his work responsibilities and listed those with whom he worked (' Role senders ').

The role senders were then interviewed to determine their preferences as to how the focal person should do his job. Finally, the focal person was reinterviewed for his own perception of how he did his job. The degree of matching of these expectations provided the measure of role conflict.

The data show that role conflict is more intense among those who are required to make contacts outside their organisation, or outside their department, and also among those who are responsible for innovation. Where jobs are of low interdependence role conflicts are often resolved, at least in the short run by avoidance and denial. Where jobs are more highly interdependent the protagonists tend either to thrash out shared problem-solutions or else to suffer acutely.

Kahn, *et al.*, contend that role conflicts are inherent in any organisation, but that they are often greater than they need be. Role clarification, they argue, can best be achieved by joint problem-solving between the focal person and his role senders ; it is often prevented by the fact that a man's boss is so much more powerful than the other role senders. In short, role conflict is fostered by autocracy ; in a more permissive organisation there may be more evident fighting but there will be less crippling role conflict. This idea will need testing out in educational as well as industrial settings.

P. B. SMITH.

MAGER, ROBERT, F. (1965). *Preparing Instructional Objectives.* Palo Alto, California : Fearon Publishers ; London : Inadcon, Ltd., pp. 62, 14s.

Since Bloom's *Taxonomy* was published in 1956, we have become increasingly aware of the need for, and the difficulty of, defining educational objectives. The spread of the use of programmed learning has emphasised this, for while in ordinary classroom teaching it is fairly easy to follow a syllabus without even defining precise goals, in programming this is the first essential step. It ought, or course, to be the first step in any curriculum planning and, according to the author of this book, the reason why it so rarely is, is because few people know now to take it. The purpose of the book is to show them how to do so. It is not concerned with making value judgments on what ought to be taught but only with the form which a usefully stated objective should take.

The book is a scrambled text divided into seven short sections. It follows a familiar pattern of exposition followed by questions to which right answers receive lavish praise and wrong ones are castigated before the student is led through a remedial sequence. The castigation is, however, sufficiently light-hearted to provoke English students to choose the wrong answers on purpose so as to see the consequences. After an introduction, the book starts by showing the difference between a *description* of a course of study such as may be found in a course handbook or examination syllabus and a statement of its *intended outcomes*. Then follows a warning of the use of equivocal terms like ' to understand,' ' to appreciate ' and ' to know ' in defining objectives and an insistence upon the value of definition in behavioural terms. Subsequent sections deal with identifying the terminal behaviour which will give evidence that the learner has reached the intended goal, to describing the conditions under which the behaviour is to occur and to defining the criterion of acceptable performance. Finally, there is a test in which the reader is told that his acceptable performance is not more than seven mistakes out of forty-four items. A group of teachers taking an advanced course in educational psychology found they could achieve this level in under two hours. Younger students would obviously take much longer, perhaps twice as long.

In level the book is appropriate for students in colleges of education, but it will also appeal to practising teachers who want to start making their own instructional programmes. Though its claims to being unique are probably over stressed, for there are many texts which deal with defining objectives, including a previous version of this one, it is a useful little book at this level. L. B. Birch.

Meyer, Wm. J. (1964). *Developmental Psychology*. New York : The Center for Applied Research in Education, Inc.

Barbe, Walter, B. (1963). *The Exceptional Child.* Washington, D.C. : The Center for Applied Research in Education, Inc., pp. ix+117, 32.

Dr. Meyer's survey of children's individual development reflects the movement away from the earlier descriptive maturational approach with its emphasis upon measurement and the establishment of norms. He takes his stand upon the interaction of genetic and environmental factors and not only insists upon the complexity of this interaction, but also reveals the range of variables in behaviour on the one hand and environmental factors on the other, which need to be related. He dictates some of the gaps in knowledge and some of the research required to fill them.

The theoretical structure he adopts offers a basis for the manipulation of the environment, the development of readiness and the changes in behaviour which will satisfy the individual differences of each developing child. This discussion of sequences of behaviour, e.g., in perception, in concept formation and manipulation, and in cognitive style produces more fruitful suggestions than can be offered by S-R theories or by ' field ' theories. The former lack developmental implications and the latter, to which he leans, because of their loose conceptual structure, do not favour the analysis of those variables with which he is concerned. Dr. Meyer's discussion is not unsympathetic to the principles underlying the work of the Manchester school.

He regards social behaviour as " a dependable indicator of the adequacy of psychological development " (p. 87). He is not unaware of the importance of parents' and teachers' influences, but tends to regard peer relationships primarily as a means of satisfying children's needs. The view of social relationships as a variable of comparable importance with those he discusses seems to be inadequate.

The standard set by D. Meyer's book in this series, The Library of Education, is not maintained by Dr. Barbe's monograph on The Exceptional Child. Its purpose is to underline the education of each exceptional child according to his needs and capabilities. Its form is prescriptive, based upon a series of generalisations some of which are out-of-date and others unsubstantiated by evidence. In one place (p. 9) he bases his argument upon the relative constancy of the I.Q., and in contrast to Dr. Meyer, he neglects the complexities of evidence and argument that have accumulated round this doctrine.

This book is not addressed to psychologists. The teacher receives little illumination or guidance in her procedures. Neither will she find a conceptual framework with which to approach the problem of exceptional children. The author's concern with classification and the dangers of improperly labelling children suggests that the book is addressed to administrators. He is, perhaps, more concerned with the attitudes of American society " which has placed such great emphasis upon conformity " (p. 97), and is intended to develop the variety of contributions of exceptional children to the democratic states. J. A. Simms.

Reeves, Joan W. (1965). *Thinking about Thinking.* London : Secker and Warburg, pp. 333, 55s.

The opening sentence of the author's introduction indicates that, " This is a book of exploratory essays on approaches to thinking," and that is the impression which remains with the reader. The plan does not consistently follow that of the source book, although some interesting data which are not always readily available have been assembled ; nor does the work consist in the systematic appraisal of particular positions or the defence of a thesis or point of view.

A variety of aspects of thinking are considered, usually from the standpoint of authors who have had something interesting and worthwhile to contribute and one of the most pleasing features is that much of the treatment, whether theoretical or empirical, deals with genuine, qualitative features in thinking. Locke's empirical approach is related to the oft-neglected psychology of his part contemporary, Spinoza, and while no attempt is made to diminish the contribution of Freud to the qualitative study of thought, an interesting review of his predecessors in the consideration of unconscious thinking precedes the treatment of Freud's own work in this field. Attractive features of the approach of Galton and particularly of Binet, are reviewed at some length, including the latter's excursions into the methods employed by prodigious calculators and famous chess players. Later contributors receive much less attention; but there are frequent observations which are provocative and valuable of which perhaps the most outstanding are those relating to Suttie's *The Origins of Love and Hate*.

There are, too, other aspects of provocation for the staid reader. The author makes frequent use of the rhetorical question and intrudes with personal apologies or viewpoints in the course of passages of direct exposition. Again, it is a matter for regret that Ernest Jones' valuable contribution on 'The Theory of Symbolism' along with Freud's observations on the development of language and the antithetical sense of common words were not woven more clearly into the treatment of Freud's considerable contributions to the nature of thinking itself. So, too, younger readers could perhaps be reminded of the existence of the Vineland publications and that Binet, despite his deservedly great reputation, had an able collaborator.

There is, nevertheless, a refreshing eclecticism associated with the choice of authors discussed and the examples cited so that even the provoked or mildly exasperated reader would be grateful for the book. F. V. SMITH.

STAATS, A. W. (1964). *Human Learning*. New York: London: Holt, Rinehart and Winston, pp. xv+520, $8.75.

This book is primarily a collection of well-nown recent papers put together by Dr. Staats into the following sections: Child Learning; Verbal Behaviour; Communication and other functions of Language; Problem Solving; Attitude Learning and Human Motivation; Social Interaction; Personality and Behavioural Treatment.

Dr. Staats has written a short introduction to each of the sections and makes some attempt to link the papers together. The main orientation of most of the contributors is in the conditioning traditions of American psychology. Much of it makes for pretty dull reading, set out in the now familiar jargon and sometimes taking a long time to say the obvious. Of course, some of the individual papers by well-known authors are important and have been recognised as such, but the total impression of the collection is far from impressive. If this is all that the psychologist has to say in the field of human learning to the educationalist, it is a disappointing contribution. Conditioning is an important principle, but in trying to attribute so much to it, we should not lose sight of how little we understand. Conditioning describes with some precision the particular circumstances in which learning takes place, and in so far as conditioning is primarily an account of the spatial-temporal relationships governing an organism, it could hardly fail to give a plausible account of those circumstances for we are discussing the major physical variables controlling any biological creature. But having described their influence via conditioning, we are little wiser how they are operating. In attempting to explain some of the more complex mental processes in terms of conditioning, it would be as well to ask how much do we understand about conditioning itself. We may grant that it enables us to control and even to predict behaviour, but that is quite another matter.

It seems to me unfortunate that a book of this kind should put so much emphasis upon its one theme when it is offering relatively little that is new. There are today many other approaches to human learning that are exciting and developing and lie outside this area. Dr. Saats's book will be mainly useful as a reference for some classical papers in this field. This must be its justification, and $8.75 cents for a collection of reprints need some justification. H. KAY.

YOUNG, MICHAEL (1965). ***Innovation and Research in Education.*** London: Routledge and Kegan Paul, pp. v+184, 25s.

The problem that confronts the writer on educational research is that he is not dealing with a tidy or clearly defined area of inquiry. The subjects of educational research are drawn from classroom studies, administration, psychology and sociology, and interpretation of the results requires some acquaintance with experimental design, survey techniques and statistics. The student is often overwhelmed by this intimidating picture and quickly abandons the effort to search for evidence on controversial topics. Any attempt, therefore, to present research findings and methods in a simple and stimulating way is to be welcomed. In this book, Dr. Young draws from a wide range of studies which enable the reader to ' get the feel, of the climate of research, and he attempts to stimulate ' the more usual teacher in the more usual classroom ' to respond to new ideas in a more positive way. What he proposes is action research in education which will take account of the techniques developed in the social sciences. The teacher must be involved and willing to plan and observe intelligently the consequences of action.

A useful chapter discusses the influence of fashion upon educational theory and practice. Educational psychology is seen as being particularly subject to changes in psychological theory, but this complex interaction deserves more thorough treatment and the judgment of Piaget's work in particular is less than fair. However, a suggestion which most educational psychologists would support is for the study of the influence of fashion itself upon the selection of topics and the judgement of researchers. A chapter on innovation in the classroom describes the climate of change with respect to curriculum reform, language teaching, programmed learning, examinations and team teaching.

The second half of the book makes tougher reading for the student. The topics are research method and the organisation of research together with some vary useful appendices on research in Institutes and Departments of Education by the Institute of Community Studies and William Taylor, and research in the U.S.S.R. and Australia by Dr. W. D. Wall and S. S. Dunn, respectively. For those actively concerned in promulgating research this is the most interesting and rewarding part of the book.

Within the scope of quite a short book Dr. Young throws up many interesting and challenging ideas. He is always a stimulating writer and the reviewer has used the book successfully as the basis of student discussions. It is in opening up thoughts about many aspects of educational research, that its chief value lies.

P. K. POPPLETON.

YOUNGHUSBAND, E. (1965). ***Social Work with Families.*** London: Allen and Unwin, Ltd., pp. 1-187, 25s.

This book of readings is one of a series of books on different aspects of social work published by the National Institute for Social Work Training. It consists of fifteen papers published on both sides of the Atlantic since 1957, in journals which are not easily consulted by the majority of social workers or teachers. The papers are arranged in three sections dealing with normal and less normal families as a group, with crisis situations for children, and with attempts to establish a theoretical framework for studying the family. The emphasis is upon an interaction approach which considers the relationships between parents, children and helpers rather than each separately.

Reading these papers is a useful reminder to those who are engaged in the relatively impersonal processes of education, that both parents and children need skilled help and guidance in stressful circumstances. They can be recommended as a source book for all who work with children. P. K. POPPLETON.

PUBLICATIONS RECEIVED

The mention of a book in this list neither implies, nor precludes, a later review.

Ruth Landes (1965). *Culture in American Education.* New York: John Wiley and Sons, pp. vi+330, 60s.

B. M. Foss (1965). *Determinants of Infant Behaviour, III.* London: Methuen and Co., pp. xiii+264, 50s.

Richard C. Atkinson, Gordon H. Bower, and Edward J. Crothers (1965). *An Introduction to Mathematical Learning Theory.* New York: John Wiley and Sons, pp. xiii+429, 75s.

Dominick A. Barbara (1965). *Questions and Answers on Stuttering.* U.S.A.: Charles C. Thomas, vii+102, $5.50.

H. B. Gibson (1965). *Manual of the Gibson Spiral Maze.* London: University of London Press, Ltd., pp. 12, Manual 4s., Card 4d.

Massialas and Smith (1965). *New Challenges in The Social Studies.* California: Wadsworth Publishing Co., Inc., pp. viii+261, 32s.

Clarke (1965). *Statistical Reasoning and Procedures.* Charles E. Merrill Books, Inc., London: Prentice-Hall International, ix+390, 56s.

Torrance (1965). *Constructive Behavior: Stress, Personality and Mental Health.* California: Wadsworth Publishing Co., Inc., Prentice-Hall International, pp. x+432, 56s.

L. Dodge Fernald (1965). *Experiments and Studies in General Psychology.* London: George G. Harrap and Co., pp. xiii+200, 15s.

Wm. H. Fitts (1965). *The Experience of Psychotherapy.* New Jersey: D. Van Nostrand Co., Inc., pp, 188, 15s. 6d.

P. Milner and S. Glickman (1965). *Cognitive Processes and the Brain.* New Jersey: D. Van Nostrand Co., Inc., pp. viii+215, 15s. 6d.

R. L. Sutherland and B. Kruger Smith (1965). *Understanding Mental Health.* New Jersey: D. Van Nostrand Co., Inc., pp. vi+233, 15s. 6d.

George A. Cicala (1965). *Animal Drives.* New Jersey: D. Van Nostrand Co., Inc., pp. vi+265, 15s. 6d.

M. L. Kellmer Pringle (1966). *Investment in Children.* London: Longmans, Green and Co., pp. xi+180, 12s. 6d.

R. R. Dale and S. Griffith (1965). *Down Stream.* London: Routledge and Kegan Paul, pp. x+97, 20s.

P. Horst (1965). *Factor Analysis of Data Matrices.* New York: Holt, Rinehart and Winston, pp. xix+730, 120s.

A. Anastasi (1965). *Individual Differences.* New York: John Wiley, pp. xiii+301, 40s. and 25s.

D. W. Harding (1966). *Social Psychology and Individual Values.* London: Hutchinson and Co., pp 184, 12s. 6d. pr.

J. Klein (1966). *Working With Groups.* London: Hutchinson and Co., pp. 240, 12s. 6d. pr.

J. Klein (1966). *Working With Groups.* London: Hutchinson and Co., pp. 240, 30s. hd.

D. W. Harding (1966). *Social Psychology and Individual Values.* London: Hutchinson and Co., pp. 184, 30s. hd.

S. Jackson (1966). *Special Education in England and Wales.* London: Oxford University Press, pp. 147, 13s. 6d.

W. R. Reitman (1965). *Cognition and Thought.* New York: John Wiley and Sons, pp. xiii+312, 60s

J. F. Perez (1965). *Counseling: Theory and Practice.* U.S.A.: Addison-Wesley, pp. v+186, 38s.

SOL. ADLER (1966). *A Clinician's Guide to Stuttering.* U.S.A.: Charles Thomas, pp. xiii+175, no price.

—— (1965). *The Excitement and Fascination of Science.* U.S.A.: Annual Reviews, Inc., pp. x+566, $5.

A. K. C. OTTAWAY (1966). *Learning through Group Experience.* London: Routledge and Kegan Paul, pp. viii+168, 25s.

C. I. SANDSTROM (1966). *The Psychology of Childhood and Adolescence.* London: Methuen and Co., pp. 259, 30s.

—— ±(1965) *L'Annee Psychologique.* Paris: Presses Universitaires de France, pp. vii+308, no price.

F. MUSGROVE (1966). *The Family, Education and Society.* London: Routledge and Kegan Paul, pp. viii+156, 21s.

D. E. M. GARDNER (1966). *Experiment and Tradition in Primary Schools.* London: Methuen and Co., pp 211, 27s. 6d.

O. PFUNGST (1965). *Clever Hans: The Horse of Mr. Von Osten.* New York: Holt, Rinehart and Winston, Inc., pp. xxxix+274, 34s.

G. M. STEPHENSON (1966). *The Development of Conscience.* London: Routledge and Kegan Paul, pp. x+134+14, 25s.

R. L. GREGORY (1966). *Eye and Brain.* London: Weidenfeld and Nicolson, and World University Library, pp. 254, 14s.

M. ROBINSON (1966). *A Basic Science Course for Secondary Schools.* London: Longmans, pp. viii+222, 13s. 6d.

CHASE and LUDLOW (1966). *Readings in Educational and Psychological Measurement.* Boston: Houghton Miffin Co., pp. xi+381, $4.25.

S. S. SLAUGHTER (1966). *The Educable Mentally Retarded Child and His Teacher.* Philadelphia: F. A. Davis and Co., pp. 191, $4.50.

From Saunders: Three Essential Books on Psychology

NEW!

Hebb: TEXTBOOK OF PSYCHOLOGY
New (2nd) Edition

By Donald Olding Hebb, Ph.D., *Professor of Psychology, McGill University, Montreal, Canada.*

The *New (2nd) Edition* of this best-selling text offers a revised and expanded clarification and codification of the ideas which make up the main structure of psychological theory. Departing radically as it does from the "popular science" approach of many introductory general psychology texts, this up-to-date basic principles text gives the student a clear picture of today's theories of the mechanisms of human behaviour.
In this revision, many topics have been expanded and enlarged. The treatment of learning has been changed and its now more extensive than before. More attention has been given to the subject of emotion and its social explanation.

353 Pages Illustrated **£2 0s. 0d.**

NEW!

Whittaker: INTRODUCTION TO PSYCHOLOGY

By James O. Whittaker, Ph.D., *Professor of Psychology, North Dakota State University, Fargo, North Dakota.*

Here is a superb new psychology text for beginning college students—carefully balanced in content and emphasis to offer an optimal amount of material for the one-semester course. The keystone of the book is its strong coverage of *motivation, learning* and *perception.* This is an integrated text with one topic logically setting the stage for the next. Current research activities are well represented and evident throughout the body of the text. Six eminent psychologists have contributed chapters—*Sir Frederic Bartlett, Muzafer and Carolyn Sherif, S. M. Luria, Russell Sergeant and Robert D. Meade.*

631 Pages. Illustrated. **£2 12s. 6d.**

Whittaker:
STUDENTS' MANUAL TO ACCOMPANY INTRODUCTION TO PSYCHOLOGY

By James O. Whittaker, Ph.D.

This manual contains nearly 400 provocative questions/problems broken down chapter by chapter to match the text. Answers are given in the appendix. Each page is perforated for easy removal and punched for a 3-ring notebook.

224 Pages. **£1 3s. 0d.**

W. B. Saunders Company Ltd. 12 DYOTT STREET
LONDON, W.C.1

THE SOCIOLOGY OF EDUCATION

P. W. Musgrave

'There should. . . be a welcome for Mr Musgrave's admirable book and it should soon figure on many reading lists in university education departments and colleges of education . . . carefully designed to meet the needs of students preparing to become teachers.' *The Times Educational Supplement.* **30s**

EXPERIMENT AND TRADITION IN PRIMARY SCHOOLS

D. E. M. Gardner

'The fruit of twelve years of dedicated research, this book compares the attitudes and attainments of children educated at 'experimental' and at more 'traditional' schools. It provides solid and fascinating evidence in support of the claim that children benefit most where the academic programme is tailored to meet their own needs.' *Western Mail.* **27s 6d**

Methuen's Manuals of Modern Psychology

THE PSYCHOLOGY OF CHILDHOOD AND ADOLESCENCE

C. I. Sandström

The psychology of childhood from the pre-natal stages up to adolescence and beyond is described in this book, which effectively fills the need for a book on developmental psychology suitable for use by students. **36s**

MOTIVATED LEARNING

A DEVELOPMENTAL STUDY FROM BIRTH TO THE SENIUM

Miriam E. Hebron

The main theories of motivation and learning as they apply to each phase of the human life-cycle are comprehensively discussed, with special attention to the period of adolescence. **35s**

METHUEN

IDENTIFICATION AND CHILD REARING

Robert R. Sears, Lucy Rau, and Richard Alpert

This research project, a major contribution to developmental psychology, explores the association between various child-rearing practices and certain behavioural developments dependent on the identification process. 63*s net*

STUDIES OF TROUBLESOME CHILDREN

D. H. Stott

A challenging re-examination of the causes of maladjustment that offers consistent evidence based on extensive case material for a common congenital factor underlying behaviour disturbance. 32*s net*

ABSENT

Max B. Clyne

SCHOOL REFUSAL AS AN EXPRESSION OF DISTURBED FAMILY RELATIONSHIPS

In this study, which is illustrated by case material from his own general practice, Dr Clyne attempts to clarify the concept of school refusal, to establish differential diagnostic criteria, and to elucidate the dynamics and aetiology of the condition.

September 38*s net*

PSYCHOTHERAPY AND CHILD DEVELOPMENT

Jean Biggar

AN INTRODUCTION FOR STUDENTS

A child psychiatrist discusses, with a wealth of clinical examples, the formative part played by the experiences of the early years in establishing the foundations of character and of mental health.

September 25*s net*

PLANNING FOR EDUCATION IN PAKISTAN

Adam Curle

A PERSONAL CASE STUDY

Professor Curle draws upon his wide experience as educational planner and adviser on social development in some of the newer nations to indicate paths to the goal of universal education, with emphasis on the primary and technical levels.

Autumn 42*s net*

TAVISTOCK PUBLICATIONS

The British Journal of
Social and Clinical Psychology

Edited by MICHAEL ARGYLE and R. R. HETHERINGTON

Contents of Vol. 5, Part 1, February, 1966

AHODA, G. Impressions of Nationalities—an Alternative to the 'Stereotype' Approach.

MORRISON, A. and McINTYRE, D. The Attitudes of Students towards International Affairs.

MURDOCH, P. H. J. Birth Order and Age at Marriage.

RIM, Y. Machiavellianism and Decisions Involving Risk.

WILDE, G. J. S. The Significance of Decided and Query Answers to Lie-Items in a Personality Questionnaire.

BARBER, T. X. The Effects of 'Hypnosis' and Motivational Suggestions on Strength and Endurance: A Critical Review of Research Studies.

FRANSELLA, FAY and ADAMS, B. An Illustration of the Use of Repertory Grid Technique in a Clinical Setting.

CLARIDGE, G. S., WAWMAN, R. J. and DAVIES, M. H. Sedation Threshold, Spiral After-effect and Overinclusion.

BOOK REVIEWS

The subscription price per volume of four parts is 90s. net. Orders may be sent to your bookseller or direct to the publishers

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200, Euston Road, London, N.W.1

THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY ARTHUR SUMMERFIELD

CONTENTS OF VOLUME 56, PART 4, 1965

NEISSER, ULRIC and BELLER, HENRY K. Searching through word lists.
NEISSER, ULRIC and STOPER, ARNOLD. Redirecting the search process.
TREISMAN, ANNE M. The effects of redundancy and familiarity on translating and repeating back a foreign and a native language.
FURTH, HANS G. and YOUNISS, JAMES. The influence of language and experience on discovery and use of logical symbols.
MATTHEWS, W. A. The relation between association norms and word frequency.
BROWN, W. P. Emotional indicators in word association.
MURDOCK, BENNET B., JR. Effects of a subsidiary task on short-term memory.
PICKFORD, R. W. Two artists with protan colour vision defects.
CORTEEN, R. S. and BLACKMAN, A. R. Skin conductance change and sensory discrimination.
MCFARLAND, JOSEPH H. Sequential part presentation: a method studying visual form perception.
WARRINGTON, ELIZABETH K. The effect of stimulus configuration on the incidence of the completion phenomenon.
HERMELIN, B and O'CONNOR, N. Visual imperception in psychotic children.
SCHWARTZ, STEWART and SPITZER, GILBERT. Preferred handedness as a factor in the ability to form a set based on imaginal situations.
MURCH, GERALD M. A simple laboratory demonstration of subception.
ORME, J. E. Ability and season of birth.
ARTHUR, A. Z. 'Cross-over' effect and subjective intensity.
AUDLEY, R. J. and WALLIS, C. P. 'Cross-over' effect; a reply to Arthur.
MCPHERSON, FRANK M. Comment on Eysenck's account of some Aberdeen studies of introversion-extraversion and eyeblink conditioning.
PUBLICATIONS REVIEWED.
OTHER PUBLICATIONS RECEIVED

Subscriptions 90s. net per volume of 4 parts
Orders may be placed with a bookseller or sent direct to the publishers

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London, N.W.1

OCCUPATIONAL PSYCHOLOGY

Editor : ALEC RODGER

OCTOBER 1965 VOLUME 39, No. 4

CONTENTS

A Study of Supervisors in the Iron and Steel Industry. By PETER B. WARR, MICHAEL W. BIRD and RAYMOND W. HADFIELD

Food, Drinks and Sweets in the Reduction of Industrial Fatigue. By ROSEMARY CASS-BEGGS and F. E. EMERY

The Union Branch in the Steel Industry : A Socio-Psychological Interpretation. By JOHN CHADWICK-JONES

Exploring Characteristics of the Versatile Worker. By D. PYM

Choice of Career by Grammar School Boys. By GRAHAM B. HILL

A Study of the Ability of Experienced and Naive Subjects to Recognise a Change in the Rate of Movement of a Lever. By E. N. CORLETT

Book Reviews

Other Books Received

Book Review Index

Annual Subscription 40 shillings, overseas 50 shillings

National Institute of Industrial Psychology

14, Welbeck Street, London, W.1.

THE BRITISH JOURNAL OF DISORDERS OF COMMUNICATION

Incorporating **Speech Pathology and Therapy**

Editor

MURIEL E. MORLEY, D.Sc., F.C.S.T., F.A.C.S.T.(Hon.).

Editorial Board

MRS. P. BRAUN	Basel
MISS M. McGOVERN	Edinburgh
MISS J. MITCHELL	Sunderland
DR. MAX NELSON	California
DR. T. PERELLO	Barcelona

The scope of the journal is disorders of communication in children and adults including the receptive and expressive aspects of language, audiology, articulation and personal, social and environmental conditions relating to communication through spoken and written language.

Publication Dates April, October.

Annual Subscription 30*s*. *Single Issues* 15*s*.

Contents of Volume 1, Part 1, April 1966

Speech Therapy comes of Age
DR. ANNE H. McALLISTER

Investigations into Speech Disturbances following Stereotaxic Surgery for Parkinsonism
MISS CHRISTINE M. ALLAN

Perceptual Disturbances in Children with Central Nervous System Dysfunctions and Implications for Language Development
DR. JON EISENSON

Hearing in Relation to Language Disorders in Children
PROF. IAN G. TAYLOR

Some Dimensions of Aphasic Impairment in Adults considered in relationship to Investigations of Language Disturbance in Children
DR. H. SCHUELL

Studies in Aphasia : a Preliminary Report of a Psycholinguistic Method for Classifying Aphasia
DR. JOSEPH WEPMAN and DR. L. V. JONES

Book Reviews

E. & S. LIVINGSTONE LTD., 15–16–17 TEVIOT PLACE, EDINBURGH

THE BRITISH JOURNAL OF MEDICAL PSYCHOLOGY

Edited by Dr. THOMAS FREEMAN

Volume 39, *Part* 2 CONTENTS *June*, 1966

DAVID R. HAWKINS : A review of Psychoanalytic Dream Theory in the Light of Recent Psychophysiologic Studies of Sleep and Dreaming.

PETER L. DIOVACCHINI : Dreams and The Creative Process.

JOHN C. DENMARK : Mental Illness and Early Profound Deafness.

J. S. MADDEN : Melancholy in Medicine and Literature : Some Historical Considerations.

H. H. GARNER. Psychotherapy and Confrontation Technique Theory.

J. S. B. LINDSAY. Probability and Schizophrenia.

ENRIQUE GUARNER : Psychodynamic Aspects of Drug Experience.

DAVID SMALL : A Multiple Choice Version of the TAT as a Measure of Aggression in Psychiatric Patients.

REVIEWS :

Subscription £4 10s. per volume of 4 parts. Single parts 30s. plus postage

Orders may be placed with any bookseller, or direct to :

CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200, Euston Road, London, N.W.1

AUSTRALIAN JOURNAL OF PSYCHOLOGY

Editor : I. K. WATERHOUSE, Macquarie University, N.S.W.

Vol. 18 *No.* 1 *April*, 1966

ROSS, J. The relationship between a Jungian Personality Inventory and Tests of Ability, Personality and Interest.

CROPLEY, A. J. and WECKOWICZ, T. E. The Dimensionality of Clinical Depression.

SUTCLIFFE, J. P., and BRISTOW, ROSEMARY A. Do Rank Order Scale Properties Remain Invariant under Changes in the Set of Scaled Stimuli ?

MANN, L. The effect of Relations between Evaluated Entities on Attitude Change.

McKINNON, T., and SINGER, G. Drive Level and Reinforcement as Variables in Paired Associate Learning of Schizophrenics.

SEGAL, BARBARA and CHAMPION, R. A. The Acquisition of Activating and Reinforcing Properties by Stimuli Associated with Food Deprivation.

SINGER, G., and DAY, R. H. The effects of Spatial Judgments on the Perceptual After-effect Resulting from Prismatically Transformed Vision.

LIVESEY, P. J. The Rat, Rabbit and Cat in the Hebb-Williams Closed Field Test of Intelligence.

ANDERSON, J. The Readability of Australian Newspapers.

Book Reviews

The Journal is published three times a year. Subscription rates 36s. (sterling), and $5·50 in U.S.A. Members of B.P.S. and A.P.A. should enquire about special subscription rates from the Secretary of their Society. Non-members should send enquiries to Melbourne University Press, Parkville, N.2., Victoria, Australia, or to its agents.

in U.K.
Cambridge University Press,
200 Euston Road,
LONDON, N.W.1

in U.S.A.
Cambridge University Press,
32 East 57th Street,
NEW YORK 10022, N.Y.

Enquiries from the Orient should be sent directly to Melbourne University Press.

Erik H. Erikson

Insight and Responsibility

Six lectures on the ethical implications of psycho-analytic insight. The lectures, which have all been expanded and revised for this volume, are :

The First Psychoanalyst
The Nature of Clinical Evidence
Identity and Uprootedness in Our Time
Human Strength and the Cycle of Generations
Psychological Reality and Historical Actuality
The Golden Rule in the Light of the New Insight 30s.

Young Man Luther

A study in psycho-analysis and history. " He has used all his brilliant gifts of interpretation on the data . . . One of the most original and stimulating studies which have been produced in the present decade."—Geoffrey Gorer, *The Listener.* 25s.

Faber & Faber 24 Russell Square London WC1

The Birmingham Printers Ltd

PRINTERS AND BOOKBINDERS

High Class
Technical Publications

42-44 HILL STREET, BIRMINGHAM 5
Telephone : MIDLAND 0196-7

VOLUME XXXVI NOVEMBER 1966 PART 3

THE PERCEPTION OF PEER PERSONALITY BY ADOLESCENT GIRLS

By A. MORRISON
(*Moray House College of Education, Edinburgh*)

AND H. J. HALLWORTH
(*Department of Education, University of Birmingham*)

SUMMARY. Previous studies of personality assessments of adolescents in secondary schools have depended largely upon the use of teachers' ratings. For the present investigation, ratings of boys and girls from two age groups were made by girl peers.

The subjects were 220 pupils in the second and fourth years of a co-educational secondary school. Four girls in each of eight classes made ratings for boys and girls separately on twenty personality traits.

A matrix of correlations between ratings was computed for each of four categories: girls' ratings of girls, and of boys, in the second and the fourth year. A principal components analysis was made for each of the four matrices of correlations, and the axes were rotated in accordance with the Varimax method. In each case, eight factors were extracted, accounting for some 70 per cent. of the variance.

Three factors are broadly comparable across the four analyses; they have been interpreted as three dimensions on which adolescent girls perceive and classify their peers. Discussion is divided into three parts: the identification and naming of the dimensions, comparisons with the dimensions consistently reported from studies of teachers' ratings, and comparisons across the present analyses in respect of the age of raters and subjects, and the sex of subjects.

I.—INTRODUCTION.

STUDIES of teachers' personality ratings of pupils in secondary schools have consistently demonstrated that teachers classify the attributes of their pupils' on two main dimensions, one descriptive of ' reliability and conscientiousness,' and the other of ' social extraversion ' (Hallworth, 1964). Apart from small but interesting differences, the traits used to define each of these dimensions appear to be independent of the educational context in which assessments are made, and of the age and sex of pupils. Few studies of peer ratings of adolescents have been carried out (Eysenck, 1960) and similar regularities have yet to be demonstrated across the various contexts in which peer group members rather than teachers are making the assessments. One recent study (Hallworth and Morrison, 1964), in which a combined analysis was made for teachers' and peer ratings, suggested that the adolescent raters were using essentially the same dimensions as the teachers, but the information it provided was limited by the fact that the analysis was made for only one age group. Furthermore, no cross sex ratings were made, so that on two variables—age and sex of peer raters—which might well be influential in determining peer assessments, the available information is still very limited.

The present study was undertaken with these latter considerations in mind. Subjects were drawn from a different educational context, ratings were obtained for two age groups which were two years apart, and the girls who made the ratings did so for boys and girls separately. However, in terms of the techniques and the list of traits used, the study was comparable to the earlier ones mentioned above.

II.—DESIGN OF THE EXPERIMENT.

The subjects were 220 boys and girls in a co-educational secondary school in a Scottish industrial burgh and were drawn from classes in the second and fourth years (13+ and 15+ age groups). All the subjects were following courses leading to the Scottish Certificate of Education.

The raters were girl members of the eight classes concerned in the study. These girls, four from each class, were chosen at random from class lists supplied in advance by the Rector of the school. Each girl rated her classmates on twenty personality traits, as shown in Table 1 below.

TABLE 1

LIST OF PERSONALITY TRAITS USED IN THE STUDY.

1. Calmness/excitability.
2. Maturity.
3. Co-operation with teachers.
4. Persistence.
5. Prefect suitability.
6. Sociability.
7. Cheerfulness.
8. Generosity.
9. Courtesy.
10. Personal appearance.
11. Confidence.
12. Games ability.
13. Interests.
14. Originality.
15. Co-operation with other pupils.
16. Leadership.
17. Gentleness/roughness.
18. Speech.
19. Loyalty.
20. Strength/weakness.

The first condition met by the choice of traits was that of comparability with traits used in previous studies of teachers' ratings ; but further to this, several traits were included on the grounds of their interest to teachers, and their possible importance to pupils themselves in discriminating between individuals. Thus, Prefect Suitability appeared likely to be of interest to teachers, whilst Personal Appearance, Loyalty, Maturity, Interests and Speech might be of importance to pupils themselves for differentiating between assessments made for the same and the opposite sex. The measure of Speech required an assessment to be made in terms of a subject's ability to express thoughts clearly and distinctly, and was framed appropriately to avoid assessment in terms of ' broadness ' of speech. Loyalty was framed in terms of the reliance that could be placed upon a classmate, especially to keep confidences, whilst Originality emphasised the possession or otherwise of good ideas of his or her her own and a willingness to initiate novel activities. The measure of Interests made a broad distinction for purposes of assessment between the classmate who engaged actively in many interests and the one who did not.

In most respects, the procedure for making ratings followed that used for teachers in previous studies, ratings being made on five-point scales, with the mean of the scale assumed to be the mean of the particular class group on each trait. Each sex group in each class was rated on one trait before proceeding to the next, raters being asked to distribute ratings in accordance with the percentages of the normal distribution. However, in one respect the usual procedure was changed : instead of providing rating instructions in written form, they were given orally, and pictorially on a blackboard. Only after this had been done, and raters' questions had been answered, was the actual rating undertaken. Subsequent examination of the completed ratings forms showed that the procedure had been clearly understood by all but two of the thirty-two raters. The results obtained from these two pupils were not consistently distributed in the recommended form and were excluded from the study.

III.—STATISTICAL ANALYSIS OF DATA.

Since the main purposes of the investigation were to compare same and cross-sex ratings for two age groups, the individual sets of ratings were divided into four categories, in order to provide four separate analyses. A matrix of correlations between ratings was computed for each of the four categories. Each matrix was then analysed by the method of principal components, and the axes were rotated in accordance with the Varimax criterion (Kaiser, 1959).

IV.—RESULTS.

Eight Varimax factors, accounting for approximately 70 per cent. of the variance, were extracted in each analysis. It is clear from the nature of the results that they differ in several respects from those obtained from studies of teachers' ratings, and that it is necessary to describe them both in terms of consistencies across the four age and sex groups, and in relation to important specific trait differences between comparable factors from the four analyses.

In general, the individual factors extracted are of more uniform variance than is the case with factors obtained from the earlier studies of teachers' ratings referred to above. Taken together, the three largest factors in each analysis account for between 33 per cent. and 38 per cent. of the variance. Further factors, of the order of 10 per cent. of the variance and, therefore, nearly as large as some factors obtained in the earlier studies, exist in one or more of the analyses, making cross matching of factors from the four analyses rather difficult. However, there is sufficient similarity in patterns of loadings on the three largest factors to justify some generalizations. These factors are shown in Table 2, and are discussed in the present article.

TABLE 2

LOADINGS ON THREE VARIMAX FACTORS FOR EACH OF FOUR ANALYSES.

Variables	Girls Rating Girls						Girls Rating Boys					
	Year 2			Year 4			Year 2			Year 4		
	I	II	III	I	II	III	I	II	III	I	II	III
1. Calmness/Excitability	10	14	20	03	-13	27	14	-04	22	11	06	51
2. Maturity	19	05	21	05	-09	21	11	08	33	02	-01	83
3. Co-op. with Teacher	07	40	20	-03	19	03	-14	84	10	-19	30	05
4. Persistence	06	13	26	21	-04	-02	15	10	12	08	01	15
5. Prefect Suitability	-07	55	48	03	34	-16	25	58	-04	12	20	73
6. Sociability	84	10	12	92	12	13	70	-04	39	50	-30	32
7. Cheerfulness	75	04	06	70	25	11	73	-07	22	74	-03	12
8. Generosity	44	26	18	64	47	14	44	22	43	85	32	12
9. Courtesy	06	71	21	23	61	07	21	74	11	24	53	15
10. Personal Appearance	26	53	19	14	15	18	33	26	45	09	40	04
11. Confidence	37	06	25	-02	-02	63	16	14	67	37	-22	15
12. Games Ability	22	01	81	13	12	25	20	-15	39	19	-06	47
13. Interests	29	05	59	46	-22	63	26	07	68	19	06	34
14. Originality	34	26	-09	07	15	75	51	28	23	20	-09	20
15. Co-op. with Others	66	21	33	51	45	27	52	34	-02	74	03	18
16. Leadership	27	36	36	29	00	61	51	07	18	40	-02	63
17. Gentleness/Roughness	06	34	19	00	70	-10	00	11	-05	13	86	-10
18. Speech	04	78	-01	15	02	22	01	23	18	-01	11	10
19. Loyalty	26	73	-06	17	84	-03	48	27	-01	07	40	14
20. Strength/Weakness	35	10	36	-04	-01	10	31	-04	09	16	-14	76
Percentage of Total Variance	13·4	14·3	10·2	11·8	11·7	10·7	13·7	10·5	9·8	13·3	8·8	15·5

(a) *Factor I* : *Social Extraversion.*

In each of the four analyses this extracts some 12 per cent. to 13 per cent. of the total variance and has the most consistent pattern. Loadings ranging from ·44 to ·92 appear for Sociability, Cheerfulness, Generosity and Co-operation with Other Pupils, with Leadership similar in consistency of appearance but having lower loadings. In its pattern of high loadings this factor closely resembles one of the two main factors obtained from studies of teachers' ratings. Although here we do not have direct evidence from measures of sociometric status or popularity, it may be identified, by its similarity to results from sociometric studies, as a main factor in peer evaluation.

(b) *Factor II* : *The Pleasant Classmate.*

The second Varimax factor obtained from the four analyses accounts for between 9 per cent. and 14 per cent. of the variance. Traits with consistent loadings are Courtesy, Loyalty, Co-operation with Teachers, Prefect Suitability and Generosity. However, it is necessary to consider loadings as low as ·20 or ·30 in order to satisfy consistency of pattern across the analyses. Clearly, high loadings for other traits are intervening in the different analyses : to mention one at this stage, a very high loading for Gentleness for fourth year girls' ratings of boys. However, all the traits with moderate to high loadings on this second factor characterise the pupil who is dependable and courteous with others, whether they be class companions or teachers. In some respects then, this factor resembles the teacher's assessment of the reliable and conscientious pupil. The high loadings for Courtesy and, in the analyses of girls, for Loyalty emphasise its evaluative nature, expecially when girls are rating other girls. The perception is of a pleasant and dependable classmate, but not one who is socially extravert or distinguished by leadership, confidence or many interests.

(c) *Factor III* : *The Non-Academic Leader.*

Factor III ranges in variance from 10 per cent. to 15 per cent. across the four analyses. Its highest loadings are on Interests, Games Ability and (except for second year boys) Leadership, with generally lower and less consistant loadings on Maturity, Calmness and Confidence. Like Factor II, its differences across the analyses are considerable, but the broad description is that of a confident pupil, able in such non-academic interests as games, and influential over others. A similar factor was identified, and similarly named, in the combined teacher and peer ratings study reported by the writers in a previous article (1964). For this reason, the term ' non-academic ' is retained, although there is in the present analyses no measure of academic ability as such. In the present study, the name is more appropriate for the fourth year analyses than for the second ; and for second year boys a more suitable name for Factor II would be ' The confident boy with wide interests.'

V.—DISCUSSION.

(a) *Girls' Ratings of Girls for the Two Age Groups.*

The first factor extracted from the two analyses for girls, descriptive of the socially extraverted girl, varies little for the two age groups with regard to the high loadings for Sociability, Cheerfulness, Co-operation with Others and Generosity. With younger girls, the assessment tends to be broader and includes moderate loadings for Confidence, Strength and Originality, while with the older age group more emphasis is placed upon Interests.

Factor II, the ' pleasant classmate,' suggests that fourth year assessments are more influenced by considerations of Gentleness, and less by Speech, Personal Appearance and Leadership. Also, the fourth-year girl who may be described as the ' pleasant classmate ' is assessed as less co-operative with teachers and less suitable as a prefect. However, in comparison with girls characterised by social extraversion or non-academic leadership she is still regarded as relatively suitable for the post of prefect.

Some of the greatest differences in assessment between the two age groups are concerned with ' non-academic leadership.' In the second year this assessment embraces a number of traits descriptive of the girl who is good at games, and with many interests, but in the fourth year the picture has changed. Three factors (of which only one, Factor III, is discussed here), each accounting for between 8 per cent. and 11 per cent. of the variance, are descriptive of such a pupil, but each emphasises a particular aspect of behaviour—Games Ability, Speech or Originality. The tendency appears to be in the direction of more specific assessments, distinguished by a particular characteristic, but having in common a concern with qualities of peer leadership.

(*b*) *Girls' Ratings of Boys for the Two Age Groups.*

Several changes appear to take place in the girls' perception of boys between the second and the fourth years, principally concerning the relationship between the assessments of the ' socially extraverted ' boy and the ' non-academic leader.' For the second year pupils Factor I has its key extraversion items of Cheerfulness and Sociability associated with Loyalty, Originality and Leadership ; thus, it combines a variety of traits which are descriptive of a pleasant, dependable and influential boy. Factor III is in several respects comparable to Factor I, but is centred around Confidence and Interests. Factor II is dominated by Co-operation with Teachers, and is in this respect unlike the second factor in the girls' analyses.

By the fourth year the girls' assessments of the boys have significantly changed. Social extraversion has become clearly distinguished from other characteristics. Furthermore, in terms of the amount of variance, the most important factor is now descriptive of the boy who is a ' non-academic leader.' This assessment contains the traits of Leadership, Games Ability and Interests, but it also has high loadings on two traits which are of little importance in the first three factors from the other analyses, namely, Maturity and Strength. Factor II has also changed : for second year boys, the emphasis is on Co-operation with Teachers, but in the fourth year this characteristic is relatively unimportant in comparison with the dominant loading on Gentleness. This trait also has a high loading on Factor II for fourth year girls.

The present analyses are of particular interest to teachers in making it possible to compare the characteristics which they associate with teacher co-operation and prefect suitability with the views held by the pupils themselves. In both analyses for girls, and in the analysis for second year boys, Prefect Suitability is a quality most closely associated with the ' pleasant classmate,' that is, with the traits of Courtesy and to a lesser extent Loyalty, Co-operation with Teachers and other Pupils, Generosity and Gentleness. In addition, girls' assessments of second year girls also associate Prefect Suitability with Games Ability on Factor III. The analysis for fourth-year boys suggests that in their case the girls take a different view of Prefect Suitability : this is now highly related to ' non-academic leadership ' as defined by Leadership and Games Ability, but with the greatest emphasis on Maturity and Strength.

Previous studies of teachers' ratings indicate that teachers associate co-operativeness on the part of pupils with high attainment, persistence and conscientiousness rather than with social extraversion qualities. In the present study, for all but the fourth year boys, Co-operation with Teachers is associated with Prefect Suitability on Factor II. However, it is only in the girls' analysis for second-year boys that Co-operation with Teachers has a high loading on any of the three factors discussed. Thus, it appears generally to be assessed somewhat independently of the main groupings of pupil characteristics, coming out on one of the smaller factors in association with Persistence and Personal Appearance.

VI.—CONCLUSION.

This study, in common with those carried out with teachers and pupils in English schools, was concerned with the assessments made by raters within particular contexts rather than with the identification of dimensions of adolescent personality. For this reason, traits were selected for their likely relevance to pupils and to the classroom situation, and subjectivity was encouraged both by using raters familiar with subjects, and by distinguishing between age and sex groups.

If we compare present results with those obtained by the writers in their study of raters in an English secondary modern school (Hallworth and Morrison, 1964) agreement is most pronounced with regard to the dimension of 'social extraversion.' Similarities can also be seen between the present assessment of the 'pleasant classmate' and the teachers' assessment of 'reliability and conscientiousness' revealed in earlier studies. An interesting feature of these comparative results is the confirmation from the fourth-year analyses of a factor of 'non-academic leadership.' This factor had been tentatively identified in the secondary modern study; here it is clearly replicated, and appears as a more important dimension of assessment, especially for fourth-year boys.

Beyond the general similarities between these and previous results lie several distinctive features. In the first place, the factor structure of the present peer ratings is markedly different from the typical structure of teachers' ratings. In previous studies the first three to four factors have usually accounted for some two-thirds of the total variance; in this study for each of the four analyses six to seven factors were required to account for similar amounts of the total variance. The tendency has been towards a more analytic pattern of assessment, with some of the factors having high loadings on two or three traits only. It is not possible to say how far this is due to the different sample of traits used in the present study.

Secondly, the results clearly indicate that the perceptions of adolescents are modified by circumstances of age and sex, and it would seem likely that these modifications are greater for peer raters than for teachers. It was anticipated that changes in the form of assessments would be likely, but it was not clear what these changes would be. Where girls are rating their own sex, individual traits appear to shift from one dimension to another as the age of the raters varies. Thus, in assessments of socially extravert classmates, older girls place more emphasis on Courtesy and Interests. Also, they appear to differentiate between ,aspects of 'non-academic leadership,' apparently making sharper distinctions between their classmates.

Effects of age are most noticeable, however, for girls' assessments of boys. In the second year, the first factor is more general in its range of characteristics than is the case in the fourth year, where social extraversion is more sharply defined. Also, in the second year, the traits of Co-operation with Teachers, Sociability and Interests are key items in defining the three dimensions, but by the fourth year other traits, Generosity, Gentleness and Maturity are outstanding. Such changes are most pronounced in regard to the third factor of ' non-academic leadership,' where younger girls are principally concerned with such features as the confidence, appearance and interests of a boy, but the older girls emphasise how grown up he is, his strength, stability and leadership.

Results suggest that further studies of peer ratings, especially of a cross sex nature, would be valuable both as a source of information on the relationship between teachers' and peer assessments and on the effects of age and sex in determining the social perceptions of adolescents.

ACKNOWLEDGEMENTS.—The writers wish to thank the Rector, members of staff and girls of the school for their generous assistance with this study.

VII.—REFERENCES.

EYSENCK, H. J. (1960). *The Structure of Human Personality.* Chapter IV. London : Methuen.

HALLWORTH, H. J. (1964). Personality ratings of adolescents : a study in a comprehensive school. *Brit J. Educ. Psychol.*, 34, 171-177.

HALLWORTH, H. J., and MORRISON, A. T. (1964). A comparison of peer and teacher personality ratings of pupils in a secondary modern school. *Brit. J. Educ. Psychol.*, 34, 285-291.

KAISER, H. F. (1959). Computer programme for Varimax rotation in factor analysis. *Educ. Psychol. Meas.*, 19, 413-420.

(*Manuscript received 2nd December*, 1965)

A COMPARATIVE STUDY OF CHILDREN'S BEHAVIOUR AT HOME AND AT SCHOOL

By SHEILA MITCHELL* AND MICHAEL SHEPHERD

(*Institute of Psychiatry, London University*)

SUMMARY. An enquiry carried out among a random sample of children attending local authority schools in Buckinghamshire made it possible to compare information provided by the parents and teachers of over 6,000 children. This revealed that deviant behaviour at home was significantly associated with lack of academic success and also with the manifestation of behavioural disorders in school. There were, however, many children who exhibited disorders of behaviour either only at school or only at home. This suggests that any comprehensive attempt to estimate the distribution of maladjustment in the child population must utilize the information of both teachers and parents.

I.—INTRODUCTION.

FOR many years there has been considerable speculation about the extent to which ' behaviour problems,' ' maladjustment,' or ' emotional disorders ' exist among the child population as a whole. Many studies designed to investigate this problem have been carried out in the school situation, including those of Olson (1930), McFie (1934), Rogers (1942), Cummings (1944), Ullman (1952), Bower (1958) and Stott (1958). All these studies imply that the existence of ' maladjusted ' or ' problem ' behaviour in school is a reflection of a general emotional disturbance which will also be apparent in other aspects of the child's life. Few studies have attempted to estimate the association between the child's behaviour in the two environments though Schonell (1952) has epitomized common knowledge with his remark that " It must always be born in mind that some children present a form of dual personality in respect to home and school." In the main, this deficiency may be due to the greater administrative difficulties involved in obtaining information from a representative group of parents and to the mistaken belief that parents would prove less co-operative than teachers and less reliable in their reports. Most workers who have obtained a great deal of information from a sample of mothers (MacFarlane, Allen and Honzik, 1954 ; Lapouse and Monk, 1958) have not obtained comparative data from schools. Those studies which have attempted to relate adjustment at home and at school show some disagreement in their findings. Douglas (1964), for example, dealing with a national sample of children, found that teachers rated as ' highly strung' a high proportion of 11-year-old children whose mothers reported them to be suffering from one or more of the following ' symptoms ' : bedwetting ; nightmares ; abdominal pain or recurrent vomiting ; nail-biting; thumbsucking or other such habits. Children with such behaviour problems were also found to show poor attainment and concentration at school. Glidewell, *et al.* (1963) in St. Louis, also found in a sample of 830 third-grade children that the number of ' symptoms ' reported by the mother showed a positive relationship to the degree of maladjustment reported by the teacher. The pilot investigation carried out in Birmingham for the Underwood Committee (Ministry of Education, 1955) found that psychologists' clinical ratings of the questionnaires completed by parents and teachers agreed in 60–65 per cent. of cases. Nevertheless, the same investigation showed that, of sixteen children later seen by a psychiatrist at their parents' request, eleven had been rated free of maladjustment on the basis of their teachers' reports. Similar disagreement is revealed in a small-scale clinical enquiry carried out by Piltzer (1952) who found that fifteen out of forty children selected as well-adjusted by teachers, did not appear so on Rorschach testing and that eight of these were found to show maladjusted behaviour at home.

---

* Now at the Department of Epidemiology and Preventive Medicine, The University of Glasgow.

II.—METHODS OF INVESTIGATION.

In 1961, questionnaires dealing with behaviour, health and family background were sent to the parents of a one-in-ten random sample of children aged between 5 and 15 years, who were attending local authority schools in Buckingshamshire: approximately 6,300 questionnaires (93 per cent.) were returned, completed. At the same time, separate questionnaires were also sent to the child's school to be completed by his, or her, class teacher. In this case, approximately 6,600 (97 per cent.) were received back. Completed questionnaires from both teachers and parents were available for 6,077 children, comprising 1,870 boys and 1,735 girls, aged 5 to 10 years, and 1,206 boys and 1,266 girls aged 11 to 15 years.

The information sought from parents about their children's behaviour was of two main types. (1) A series of twenty-two triple choice questions on behaviour traits as measured by the parent's estimate of *intensity* of behaviour (extreme, moderate, absent), e.g., afraid of the dark when in bed at night—a little uneasy without a light—not at all afraid of the dark. Of these items, fourteen were adapted from the fifty-three triple-choice items used by Cattell and Coan (1957) in their study of personality variables. (2) Fifteen items in which the parent had to record an impression of the *frequency* of certain kinds of behaviour (for instance, crying, nailbiting, headaches) on an eight-point scale ranging from " Never or less than once a year " to " Every day or nearly every day."

The items of behaviour included are described in detail elsewhere (Mitchell, 1965). They covered most of the behaviour problems described in the standard works on child psychiatry as occurring in clinic-attending, and therefore, presumably maladjusted, children of school age. The items *excluded* were: (i) the more bizarre forms of abnormal behaviour, e.g., autism, hallucinations, obsessional behaviour ; (ii) sexual problems ; and (iii) certain kinds of blatantly delinquent behaviour like arson. Such types of behaviour appeared to be rare even in child guidance samples and it was felt that the inclusion of questions about 'wicked' or 'mad' behaviour might worry or antagonize parents and so diminish the response rate. All the items included in the form had previously been piloted for intelligibility and lack of ambiguity by interview with mothers throughout the county.

The teachers' questionnaire covered attendance, attainment and physical disabilities as well as the presence of behaviour ' problems ' in school.

In attempting to relate the children's behaviour at home to other aspects of their life, we felt it desirable to obtain some overall index of disturbance. Originally, it was intended to use for this purpose the child's total problem score, obtained by adding together the total number of extreme items underlined by the mother. This crude score would have given each ' problem ' the same weight without regard to the frequency with which it occurred. Information gained in the survey about the distribution of each behaviour trait showed clearly, however, that some types of ' problem ' behaviour were very much more common than others. For instance, about 20 per cent. of girls were described by their mothers as having food fads. Where behaviour was reported as occurring at such a high frequency, it was felt that it was scarcely justifiable to consider it as contributing to an index of abnormality. Furthermore, it was felt necessary to make some adjustment for the changing incidence of various types of behaviour at different ages : for instance, 25 per cent. of 6-year-old boys were found to cry at least two or three times a week compared with 1 per cent. of those aged 14 years. Clearly, therefore, such crying should begiven different weight in the two age groups.

In order to allow for these factors, it was decided to adopt a concept of ' deviant behaviour ' framed in terms of a 10 per cent. limit. If any type of behaviour was reported to have occurred at an intensity or frequency found in 10 per cent. or less of any age (yearly) or sex group, then it was considered as deviant for that age and for that sex. If it occurred in more than 10 per cent. of children of that age and sex then it was discounted. Thus, in the example already given, crying two or three times a week would be considered as deviant behaviour in a 14-year-old boy, but not in a 6-year-old. An overall index of ' deviance ' was then constructed by summing the number of deviant items scored for each child. This score, then, indicated the number of types of behaviour shown by any child which were a typical of his age and sex group.

An analysis of the teachers' questionnaires revealed that only one type of behaviour, " Very quiet or withdrawn " among girls, was recorded as occurring in more than 10 per cent. of any age-sex group. The concept of deviance was not, therefore, applied to these items and the overall index of disturbance in school has merely been taken as the number of behavioural items underlined by the teacher in the questionnaire (see Appendix for list of items included).

III.—RESULTS

TABLE 1

THE RELATIONSHIP BETWEEN THE NUMBER OF DEVIANT BEHAVIOUR ITEMS UNDERLINED BY PARENTS AND THE NUMBER OF BEHAVIOUR PROBLEMS UNDERLINED BY TEACHERS.

Number of Problems Underlined by Teachers		Number of Deviant Items Underlined by Parents.							
		Boys				Girls			
		None	1–3	4–6	7–	None	1–3	4–6	7–
None	(No.)	740	846	108	35	817	847	135	47
	(%)	62%	55%	41%	43%	65%	61%	53%	45%
1	(No.)	259	387	62	20	283	308	68	29
	(%)	22%	25%	24%	24%	22%	22%	27%	28%
2	(No.)	115	158	43	10	107	127	24	13
	(%)	9%	10%	16%	12%	8%	9%	9%	12%
3 or more	(No.)	85	142	49	17	55	98	27	16
	(%)	7%	9%	19%	21%	4%	7%	11%	15%
Total (No.)		1,199	1,533	262	82	1,262	1,380	254	105
Significance of Association		Chi-squared=73 : 9 degrees of freedom : p<0·001.				Chi-squared=46 : 9 degrees of freedom : p<0·001.			

Table 1 shows that there was a significant degree of association between the presence of deviant behaviour items at home and the reporting of problems of behaviour in school. Thus, children whose parents reported that they were free of deviant behaviour were also the most problem-free group at school. Similarly, at the other end of the scale, children whose parents reported many deviant items were three times as likely to have multiple (three or more) problems recorded at school as were those who were deviance-free at home. This relationship, however, was by no means comprehensive. Thus, Table 1 demonstrates that more than a third of the children who were reported free of deviant behaviour at home, nevertheless, exhibited at least one problem in school and that nearly half of the children whose parents underlined seven or more deviant traits were apparently problem-free at school.

As for specific types of behaviour at school, there was no overall pattern which distinguished children with many deviant items underlined by their parents from those with few. The school problems of those children who were markedly deviant at home varied with age and sex.

Among the boys at all ages, those with four or more items of deviant behaviour marked by their parents were significantly* more likely than the remainder to have been assessed by their teachers as very easily frightened (11 per cent. of deviants : 5 per cent. of others aged 5 to 10 years ; 5 per cent. of deviants : 2 per cent. of others aged 11 to 15 years), to have been found disinterested in school work (12 per cent. : 6 per cent. at 5 to 10 years ; 17 per cent. : 5 per cent. at 11 to 15 years) and to have been recorded as having stolen on one or more occasion (4 per cent. : 1 per cent. for all ages). Among boys in the 5 to 10 years age group, but not among the older boys, those with four or more deviant items marked by their parents were significantly more likely to have been reported as sucking their thumbs (7 per cent. : 3 per cent.), wetting of soiling themselves at school during the current school year (5 per cent. : 1 per cent.) and as being quiet and withdrawn (17 per cent.: 10 per cent.). Among the boys of 11 or older, on the other hand, the deviant group were more likely than their fellows, to have been marked as unco-operative in class (9 per cent. : 2 per cent.), often telling lies (12 per cent. :1 per cent.), very irritable (8 per cent. : 3 per cent), aggressive towards other children (8 per cent. : 3 per cent.) and worrying nore than other children (8 per cent. : 3 per cent.).

Among the girls the only teachers' item which differentiated the deviant group at all ages was "Often tells lies" (4 per cent. : 1 per cent. in each age group). Among the older girls, those with four or more deviant items marked by their parents were also more likely to have been described as "Not interested in school work" (12 per cent. : 4 per cent.) but this was the only item which distinguished them from the rest of the girls in their age group. Among the younger girls, on the other hand, the deviant children were more likely to be reported as very restless (11 per cent. : 4 per cent.), very irritable (6 per cent. : 3 per cent.) and easily frightened (9 per cent. : 5 per cent.) in school.

Academic success in school was found to be negatively associated with deviant behaviour in the home. Table 2 shows that the chance of achieving an above-average rating in general attainment at school was more than five times as great for boys, and four times as great for girls, who were free of deviant behaviour at home as it was for those who showed seven or more deviant items of behaviour, and about twice as great as for children with four or more items noted by their parents.

---

* All the differences described were significant at least at the 5 per cent. level, using the Chi-squared test with Yates' correction for continuity.

TABLE 2

The Relationship Between the Number of Deviant Behaviour Items Underlined By Parents and the Child's General Attainment Level at School.

General attainment level relative to others of the same age at the same type of school.	Number of Deviant Items Underlined by Parents.							
	Boys				Girls			
	None	1–3	4–6	7–	None	1–3	4–6	7–
Above Average (No.)	264	279	25	3	335	279	36	7
(%)	23%	19%	10%	4%	30%	21%	15%	7%
Average (No.)	565	684	110	38	649	715	127	58
(%)	49%	46%	43%	49%	58%	53%	52%	58%
Below Average (No)	331	522	122	37	132	342	80	35
(%)	28%	35%	47%	47%	12%	26%	33%	35%
Significance of Association	Chi-squared=40 : 6 degrees of freedom : p<0·001				Chi-squared=85 : 6 degrees of freedom : p<0·001.			

TABLE 3

The Relationship Between The Number of Deviant Behaviour Items Underlined By Parents and the Child's Usual Position in the Immediate Teaching Group.

Number of deviant Behaviour Items Marked by Parent	Position in the Teaching Group					
	Boys			Girls		
	In top 25%	In middle 50%	In bottom 25%	In top 25%	In middle 50%	In bottom 25%
None (No.)	352	555	250	441	566	220
(%)	43%	40%	32%	48%	41%	37%
1–3 (No.)	405	688	400	402	657	276
(%)	50%	49%	51%	43%	47%	47%
4–6 (No.)	41	123	95	68	116	65
(%)	5%.	9%	12%	7%	8%	11%
7 or more (No)	13	33	35	17	56	30
(%)	2%	2%	4%	2%	4%	5%
Significance of Association ..	Chi-squared=40 : 6 degrees of freedom : p<0·001.			Chi-squared=42 · 6 degrees of freedom : p—0·001.		

A similar pattern emerged with respect to the children's position in their immediate teaching group (see Table 3) where the proportion of children with four or more items of deviant behaviour marked by their parents was about twice as great among those in the bottom quarter of the class as among those near the top.

IV.—Discussion.

The relationship between poor school attainment and the presence of behaviour disorders in the school setting has been discussed in several previous studies, including Burt (1937), Stott (1958) and Chazan (1958). The present study indicates that such a relationship also exists between behaviour at home and school conduct and attainment. From the information available, however, is is not possible to postulate any causal relationship. Burt and Howard (1952) found that maladjustment, defined in terms of referral for psychological investigation, was correlated with the child's exposure to too high a standard of work or to contact with an uncongenial teacher among other factors. On the other hand, Pearson (1952) suggested that emotional strains rooted in the home were often the cause of poor performance and lack of concentration at school. It is also possible that some underlying factor in the child, such as slow maturation, slight brain damage or neurological defect, may affect performance both at home and at school without there being any corresponding deficiency in either environment. Unravelling such complex interactions would require more detailed personal investigation than was possible in a large-scale statistical enquiry of the present kind. The types of behaviour which characterized the deviant children in their school setting, however, support the notion of emotional rather than organic causation. Thus, in three out of the four age-sex groups, the children who were most deviant at home were distinguished from the others by a greater tendency to be uninterested in school work, to be easily frightened and to tell lies.

It should also be emphasized that though there is a significant relationship between deviant behaviour at home and the existence of problems of behaviour in school there is still considerable divergence between the two areas of conduct. Thus, the teachers' questionnaire identified 10 per cent. of children as exhibiting three or more behaviour problems in the school setting; the same proportion had four or more deviant items recorded on the parents' questionnaire. It might be considered justifiable to pick either of these groups of children as consisting of those who were most likely to be maladjusted. Unfortunately, however, the individuals picked by the two methods are not the same: only about one child in five picked as in the worst 10 per cent. by one questionnaire was also picked by the other. This fact is important in any attempt to screen the child population to obtain a general estimate of the incidence of maladjustment. At present it is impossible to say how far the difference lies in the child's actual behaviour and how far in variation in the powers of observation and frames of reference of the persons reporting the behaviour. Clearly, however, it seems very necessary that any comprehensive screening device must study the child in both of his principal environments, the home and the school.

ACKNOWLEDGEMENTS.—This study was made possible by the generous financial support of the Nuffield Provincial Hospitals Trust. The authors also wish to acknowledge the great assistance given by the Medical Officer of Health for Buckinghamshire, the Chief Education Officer, Divisional Education and Medical Officers, and Head Teachers and staff of all the schools concerned.

V.—REFERENCES.

BOWER, E. M. (1958). A process for early identification of emotionally disturbed children. *Bulletin of California State Department of Education*, 27, 1-65.

BURT, C. (1937). *The Backward Child.* London: University of London Press.

BURT, C., and HOWARD, M. (1952). Nature and causes of maladjustment among children of school age. *Brit. J. Psychol. (Statistical Section)*, 5, 39-59.

CATTELL, R. B., and COAN, R. W. (1957). Personality factors in middle childhood as revealed in parents' ratings. *Child Develop.*, 28, 439-458.

CHAZAN, M. (1963). Maladjustment, attainment and sociometric status. *University College of Swansea, Faculty of Education Journal*, pp. 4-7.

CUMMINGS, J. D. (1944). Emotional symptoms in school children. *Brit. J. Educ Psychol.* 14, 151-161.

DOUGLAS, J. W. B. (1964). *The Home and the School.* London MacGibbon and Kee.

GLIDEWELL, J. C., DOMKE, H R., and KANTOR, M. B. (1963). Screening in schools for behaviour disorders: use of mothers' reports of symptoms. *J. Educ. Res.*, 56, 508-515.

LAPOUSE, REMA and MONK, MARY (1958). An epidemiological study of behaviour characteristics in children. *Amer. J. Publ. Hlth.*, 48, 1134-1144.

MACFARLANE, JEAN W., ALLEN, LUCILLE, and HONZIK, MARJORIE, P. (1954). *A Developmental Study of the Behaviour Problems of Normal Children.* Berkeley : University of California Press.

MCFIE, B. S. (1934). Behaviour and personality difficulties in school children. *Brit. J. Educ. Psychol.*, 4, 30-46.

MINISTRY OF EDUCATION (1955). *Report of the Committee on Maladjusted Children.* London : H.M.S.O.

MITCHELL, SHEILA (1965). *A Study of the Mental Health of School Children in an English County.* Unpublished Ph D. Thesis, University of London.

OLSON, W. C. (1930). *Problem Tendencies in Children : A Method for Measurement and Description.* Minneapolis : University of Minnesota Press.

PEARSON, G. H. J. (1952). A survey of learning difficulties in children. *Psycholanal. Stud. Child*, 7, 322-386.

PILZER, E. (1952). Disturbed children who make good school adjustment. *Smith Coll. Stud. Soc. Wk.*, 22, 193-210.

ROGERS, C. R. (1942). Mental health findings in three elementary schools. *Educ. Res. Bull.* (Ohio State University), 21, 69-79.

SCHONELL, F. J. (1952). The development of educational research in Great Britain : Part 7. Maladjusted children. *Brit. J. Educ. Psycol.*, 22, 30-39.

STOTT, D. H. (1958). *The Social Adjustment of Children.* London : London University Press.

ULLMAN, C. A. (1952). Identification of maladjusted school children. *Publ. Hlth. Monogr.* (U.S.) No. 7.

(*Manuscript received* 21*st October*, 1965)

VI.—APPENDIX.

BEHAVIOUR ITEMS INCLUDED IN THE TEACHERS' QUESTIONNAIRE.

Please underline, in the list below, any items which describe this child as he often is :

(*a*) Very restless, can't sit still for a moment.
(*b*) Cries more than most children.
(*c*) Has a stammer.
(*d*) Has other speech difficulty.
(*e*) Often tells lies.
(*f*) Has stolen things on one or more occasions.
(*g*) Is very easily frightened.
(*h*) Bites finger nails.
(*i*) Sucks thumb or finger.
(*j*) Very irritable, easily becomes cross or annoyed.
(*k*) Has had one or more temper tantrums at school during this year.
(*l*) Is unco-operative in class.
(*m*) Very shy, finds it difficult to mix with other children.
(*n*) Has wet or soiled self at school during last year.
(*o*) Has noticeable twitch of face or body.
(*p*) Worries more than other children.
(*q*) Aggressive towards other children.
(*r*) Very quiet or withdrawn.
(*s*) Very moody—on top of world one minute, down the next.
(*t*) Not interested in school work.
(*u*) Has on one or more occasions during this school year shown fear of school—that is, tears on arrival or refusal to come into the building.

THE EFFECT OF LECTURE TIME ON LEARNING

By P. J. HOLLOWAY
(*Turner Dental School, Bridgeford Street, Manchester,* 15)

SUMMARY. Dental students were lectured alternately at the beginning and end of the day throughout a lecture series at which summary hand-outs were available. They were tested by a multiple-choice examination paper, and the results analysed statistically. There was no reliable evidence that the students' performance was better after attendance at the morning than at the afternoon lectures. However, they recorded significantly more correct responses to questions set on the lectures they attended than on those they missed.

As a group they showed no preference for the late rather than early lectures, although some students favoured the former while others favoured the latter.

I.—INTRODUCTION.

RESEARCH into educational methods is becoming more the province of the full-time investigator. However, there are still a great many questions that can be answered by any teacher provided that he is sufficiently interested to formulate an hypothesis and to organise his routine teaching programme so that an unbiased test can be made.

A formal teaching programme for undergraduate dental students gave an opportunity of this type when one class was scheduled to receive a series of lectures early in the morning, while another class was to attend late in the afternoon. The dental undergraduate course, like many other science courses, consists mainly of a complex mixture of practical classes, demonstrations and lectures, and it is often difficult to know how best to arrange these. In this regard, it has been a point of some discussion that a student may be more perceptive early in the day than after a period of concentrated practical work. There are those who suggest that formal lecturing should take place first thing in the morning while students and teachers are still fresh. The opportunity was taken to test the hypothesis that the time of day at which a lecture is given has no appreciable effect on the learning process. During the study, the secondary hypothesis, that lectures themselves play no useful part in the learning process, was also examined.

II.—METHOD.

Two classes of undergraduate students attended for formal lectures. They were composed of first and third year clinical students comprising fifty-three subjects in all. In order to test the hypothesis, the lecture times were re-arranged. Instead of one class attending in the late afternoon and the other first thing in the morning, as was originally intended, each class was divided into two groups. Each group then attended alternately at 9.0 a.m. and 4.30 p.m. throughout the total course of eight lectures. Thus, each student was expected to attend four morning lectures and four afternoon lectures. His colleagues in the other group attended lectures on the same topics, but in the morning instead of the afternoon and *vice versa*. In this way, each student could form his own control; his perception at the morning lectures could be compared with his perception at the afternoon lectures. The performance of his colleagues in the other group overcame the problem of differing levels of difficulty between subjects affecting the results. Attendance was not compulsory, but a roll was taken at each lecture.

Three members of staff were involved in the lecturing in an attempt to overcome the possible idiosyncracies of one lecturer. The morning and afternoon lectures on any subject were given by the same lecturer to avoid the differing abilities of staff affecting the eventual comparisons.

At each lecture, hand-outs summarising the subject material were given to the students and were available to the non-attenders if they approached the departmental secretary subsequently.

At the end of the lecture course, the students completed a compulsory thirty-question, multiple choice examination paper. This method of examination was used because it allowed for objective marking, and it also permitted questions to be asked specifically on individual lectures, without the chance of overlapping into other lecture subjects. The paper was designed according to the format suggested by Hubbard and Clemans (1961). Four types of question were set. The first consisted of completing incomplete statements from a given choice, and the second was a more intricate form of this where the answer could be more than one of the possible alternatives. The third was a test of reasoning while the fourth related to case histories. All questions had a probability of 0·2 of being answered correctly by a guess, and all had to be answered. No time limit was set, but all the students completed the paper within 30 minutes. All the questions had been set previously to other groups of students (Heffer, *et al.*, 1965), and had been corrected for ambiguities.

Because of the experimental design, there were two separate methods of analysing the results. First, each student could act as his own control. There were three possible alternatives; he attended some lectures in the morning, some in the evening and others he failed to attend. Thus, each student's results on morning classes could be compared with his own performance on afternoon classes. In the same way, his rating on lectures he did attend could be directly compared with his own performance on lectures he did not attend. By making comparisons within students in this way, the conclusions could not be biased by comparing the results of keen with disinterested students, as would have been possible if the comparisons had been made between subjects.

The second method of analysis was by question rather than by student. Each examination question was answered by some students who had attended the appropriate subject in the morning, by some who had been lectured to in the afternoon, and by some who had not attended the lecture at all.

In both methods, non-parametric statistical tests seemed most appropriate (Siegel, 1956). The number of correct answers was expressed as a percentage of the total number of questions, and the percentages compared. For example, the percentage of questions a student answered correctly on lectures he attended in the morning, was compared with the percentage of questions he answered correctly on lectures he attended in the evening.

III.—RESULTS.

The percentages of correct responses for the morning, afternoon and non-attendance groups was first compared using the Friedman two-way analysis of variance. Whether comparisons were made by students or by questions, the groups clearly differed ($p<0{\cdot}01$). Examination of the rank totals (Rj values) indicated that the best results were obtained with the morning lectures, and the worst with the lectures not attended. The groups were then examined more closely to establish the probability levels of the various differences.

The percentage of correct answers that each student obtained from morning and afternoon lectures was compared. The median percentage of correct answers for the morning lectures was 67 per cent. (mean, 65 per cent), while that for the afternoon lectures was 61 per cent. (mean, 61 per cent.). This difference was tested using the Wilcoxan matched-pairs signed-ranks test, but no significance could be established. This particular comparison involved forty-eight subjects, and so constituted a test of considerable power.

The second comparison made was between the pooled results of the lectures attended, and the results of those not attended. Here the median percentage of correct responses obtained on lectures was 62 per cent. (mean 62 per cent.), while that for lectures not attended was 50 per cent. (mean 50 per cent.). In this case, the Wilcoxan test clearly demonstrated a significantly better performance from the student when he attended the lecture than when he did not, at a probability level of less than 0·01.

When these same results were compared by questions, the differences confirmed those already established on a ' by student ' basis with the same levels of significance.

In addition, it seemed pertinent to discover whether late in the afternoon was a more popular lecture time than first thing in the morning as far as the students were concerned. By inspection of the attendance sheet it was quite clear that some students favoured one time while others favoured the other. For each student a decision was made as to whether he attended a greater proportion of morning than afternoon lectures. Thirty students attended more afternoon lectures while twenty-one attended more morning lectures. This difference is insignificant.

IV.—DISCUSSION.

The design of the experiment enabled several questions to be answered. It showed clearly that the course of instruction given did help the learning process, and was, therefore, of some value. However, if any benefit was gained at all from lectures in the morning rather than last thing in the afternoon, it was only slight. In addition, it seemed of little consequence to the students as a group when the lectures were held, although individual preferences were obvious. These conclusions may have been anticipated before the test began, but nevertheless, they enabled future teaching organisation in the department to be based on fact rather than conjecture.

The conditions under which the study was conducted tended to detract from the ultimate result. The lectures were given in a dental clinic to groups of about a dozen in an atmosphere that was not conducive to comfort, nor undivided attention, and visual aids, other than a blackboard, were not available. The subjects were of a practical nature, and were reinforced during clinical practice by the students, whether they attended the lectures or not. In addition, the hand-outs were easily available to everyone, and many students who had not attended a particular lecture subsequently asked the departmental secretary for a copy. All the test questions were based on the hand-outs, so that a detailed knowledge of these would have sufficed to have scored a very high mark. Despite this, the lectures seemed to have an effect. However, it might also be argued, that attendance at a lecture ensured that the student received a hand-out, whereas failure to attend did not guarantee this. Thus, the notes may have been more easily at hand following a lecture attended, and this may have affected the final result. Equally, it may be thought that a student not attending a lecture is more eager to study notes when faced with an examination (Joyce and Weatherall, 1957). A detailed analysis of the outcome of such a complex situation can lead only to conjecture. All that is reasonably certain is that attendance at the lectures produced a more favourable response.

This type of study has two advantages. Firstly, it tends to produce a reliable teaching programme upon which can be based some reasonable assumptions, and secondly, it encourages the teaching staff to take a lively interest in the development of their teaching methods. Each member of a small teaching department was involved, and each was equally interested in the outcome. It may be that participation in this type of work adds to the enjoyment and, therefore, to the value of one's teaching efforts. It cannot be denied that the interest produced by the experimental situation may, in itself, have heightened the effect of the lectures.

Multiple-choice examining techniques allowed a simple, objective, estimation to be made of the student's factual knowledge, and the use of non-parametric statistical methods enabled this type of data to be analysed with some degree of confidence. Both the examining technique and the method of analysis are simple to use after some study and practice, and are to be recommended for this type of investigation.

ACKNOWLEDGEMENTS.—The author wishes to thank Professor J. L. Hardwick in whose department this investigation was carried out.

V.—REFERENCES.

HEFFER, P., HOLLOWAY, P. J., ROSE, J. S., and SWALLOW, J. N., 1965. An investigation into dental undergraduate examining techniques. *Brit. Dent. J.*, 118, 334-338.

HUBBARD, J. P., and CLEMANS, W. V., 1961. *Multiple-choice Examinations in Medicine.* Philadelphia · Lea and Febiger.

JOYCE, C. R. B., and WEATHERALL, M. (1957). Controlled experiments in teaching. *Lancet* II, 402-406.

SIEGEL, S. (1956). *Non-Parametric Statistics.* New York : McGraw-Hill.

(*Manuscript received* 21*st October*, 1965)

CREATIVITY AND INTELLIGENCE*

By A. J. CROPLEY
(*University of New England*)

SUMMARY. A battery of thirteen tests, six ' convergent ' and seven ' divergent,' was administered to 320 Grade Seven children and the ' divergent ' tests scored for originality. Correlations were then subjected to principal axis factor analysis. The results indicated that the two largest factors were defined by ' convergent ' tests and ' divergent ' tests, respectively. These factors were subjected to both orthogonal and oblique rotations and it was found that, although the ' divergent ' tests indicated a separate factor, creativity, it showed a significant correlation with the factor defined by the more normal tests of intelligence, amounting in this study to ·514.

I.—INTRODUCTION.

IN recent years it has frequently been proposed that there is a separate aspect of intellectual functioning not sampled by conventional intelligence tests (Guilford, 1950, 1960; Getzels and Jackson, 1962; Torrance, 1963), and this notion has been supported by the fact that conventional tests may account for as little as one-quarter of the variance of measures like school achievement (Getzels and Jackson, 1962, p. 2). Guilford (1959) has suggested that this is because conventional tests are heavily saturated with ' convergent thinking ' skills, and take little account of ' divergent thinking.'

The factor thus postulated is defined by tests which elicit a divergent type of thinking, i.e., by tests involving the capacity to invent new responses rather than merely to reproduce those previously learned, and is usually labelled ' creativity.' It is sometimes held to be independent of the intellective mode elicited by ' convergent' tests, and its importance has been emphasized by many authors. It has been variously described as lending dignity to the human being in a computer-dominated age (Bruner, 1962), as a vital aspect of human resources (Toynbee, 1962), and as an important element in mental health (Barron, 1963).

Various authors have argued that creativity is related to unique cognitive factors (Guilford, 1950; Lowenfeld, 1958), and that it is dependent upon certain personality characteristics (Freud, 1910; Barron, 1955; Couch and Keniston, 1960; McGuire, 1961; Drevdahl, 1963; Cropley, 1965a). However, writers differ about the degree of complexity of the processes covered by the term. Hence, although Gordon (1961) and Koestler (1964) have argued that it is a unitary trait, Ghiselin (1959) reported the existence of a factor of verbal creativity and speculated that an analogous non-verbal factor might exist, Anderson (1964) obtained two factors of verbal and non-verbal creativity, and Taylor (1964) has suggested that there are at least two types of creativity. Other workers, too (McGuire, 1961; Sultan, 1962) have reported finding differing numbers of factors of creativity.

However, despite the many studies dealing with creativity as either a dependent or an independent variable, certain authors have questioned whether there is a separate dimension of creative ability at all (Burt, 1962, 1964; McNemar, 1964). Several studies report a significant correlation between creativity and conventional intelligence; Getzels and Jackson (1962, p. 5), for example, quote correlations ranging from ·02 (Andrews, 1930) to ·27 (Welsh, 1946), while they themselves obtained correlations ranging from ·131 to ·378 (1962, p. 20). Meer and Stein (1955), Ripple and May (1962), Richards, Cline and Needham (1964), and Cropley (1965b) have all reported similar findings, while Ketcham and Kheiralla (1962) compared scores on the W.I.S.C., the Stanford-Binet and the P.M.A. test with scores on a battery of creativity tests and, out of 64 correlations, obtained 54 which were significant beyond the ·05 level. Finally, some authors (Taylor, 1964 ; Vernon, 1964 ; Yamamoto, 1965) have suggested that creativity and intelligence may become independent only after some critical I.Q. level has been exceeded.

---

* This study was conducted under the supervision of C. C. Anderson, while the author was at the University of Alberta, Edmonton, Canada.

Hence, the question arises whether a specific factor is needed to account for the processes involved in solving 'divergent' problems, or whether, in the tradition of Galton (1883), creative productions result largely from the operation of 'general ability.' Burt (1962, 1964) has supported the latter point of view, arguing that, underlying tests of both convergent and divergent thinking, there is a large general factor, supplemented by smaller group factors which are distinct from each other but may occasionally overlap in tests of a mixed type. Thus, opinions on this issue include two quite incompatible views : first that 'divergent' tests do not cluster separately from 'convergent' measures, and secondly, that they do.

A number of factor studies have attempted to clarify this issue. Thorndike (1963a) and Marsh (1964) have both factorized the correlation table reported by Getzels and Jackson (1962, p. 20), and neither was able to obtain a separate factor which was defined by the creativity tests and independent of I.Q. Thorndike did obtain a factor on which the creativity tests loaded highly, but a substantial portion of this factor's variance was accounted for by the conventional test. He concluded (Thorndike, 1963b) that, although creativity tests do tend to intercorrelate significantly, there is little evidence that they imply a separate, clearly-defined factor of intellect.

The source of this disagreement lies partly in the doubtful validity and reliability of creativity measures ; unstable measures do not yield stable factor structures. The label 'creativity' implies value judgments, in which responses are required to meet professional or aesthetic criteria (Maltzmann, 1960 ; Mednick, 1962 ; Taylor, 1964) and, in any case, as Vernon (1964) has pointed out, a study analogous to Terman's (1925) intelligence study would be needed to establish the predictive validity of creativity tests. Furthermore, both Mackler (1962) and Wodtke (1964) have shown that creativity tests have low reliabilities. Consequently, for the purposes of the present study, the more 'objective' concept of 'originality' was preferred to creativity ; and in calculating 'divergent' scores, originality was operationally defined as the tendency to produce responses statistically uncommon by comparison with the responses of a criterion group. This concept was preferred because it eliminates problems associated with the doubtful validity and reliability of creativity tests, while still permitting concentration on responses which represent productions manufactured by the individual rather than merely reproduced.

The study was concerned with the following questions : is a special factor needed to account for results obtained with 'divergent' tests, which requires invention of uncommon answers rather than the reproduction of what has been already learned ? To what extent is any such factor independent of the factor underlying tests of 'convergent' processes ?

II.—METHOD.

Sample.

The sample consisted of the entire Grade Seven population of a large metropolitan junior high school in the city of Edmonton, Alberta, Canada. Of 354 children enrolled, 320 completed the entire test battery, and these 320 children comprised the sample. Mean age was 13 years 6 months (SD=16·9 months; range—11 years 1 month to 16 years 1 month), and they had a mean verbal I.Q. on the appropriate Lorge-Thorndike scale of 114·3 (S.D.=14·5; range—79 to 150). The sample included 170 boys, for whom the mean verbal I.Q. was 114·7 (S.D.=15·2; range—79 to 148), while the corresponding figure for the 150 girls was 113·8 (S.D.=13·6; range—88 to 150). The mean age of the boys was 13 years 10 months (S.D.=16·6 months; range—11 years 1 month to 16 years 1 month), and of the girls 13 years 2 months (S.D.=17·3 months; range—11 years 7 months to 16 years 1 month).

Tests.

Tests administered fell into two groups: conventional tests of ' convergent ' processes, which required the children to find a single correct answer to each problem, and tests of ' divergent ' processes which required them to produce a response for themselves.

The first group of scores included:

(1) Verbal I.Q. (Lorge-Thorndike scale, Level 4, Form A).
(2) Performance I.Q. (Lorge-Thorndike scale, Level 4, non-verbal Form A).
(3) Vocabulary (French, 1963).
(4) Inferences (French, 1963).
(5) Length Estimation (French, 1963).
(6) Academic Average.

The second group of tests included:

(1) Seeing Problems (French, 1963).
(2) Tin Can Uses (Torrance, 1962, p. 44).
(3) Consequences (French, 1963).
(4) Symbol Production (French, 1963).
(5) Circles (Torrance, 1962).
(6) Hidden Figures (French, 1963).
(7) Associations (Mednick, 1962).

Scoring of ' convergent ' tests was carried out according to the published specifications. For the ' divergent ' tests, scoring was, wherever possible, for originality; the method adopted involved allotting weights for each child's response to each item of each test, the highest weights being given to those responses which were least common within the entire sample. The response must, of course, be task appropriate and not merely bizarre. Modifying the weights suggested by Torrance (1962, p. 72) in order to increase the variance of the total scores, the following weights were employed: responses appearing on more than 15 per cent. of protocols—0; responses appearing between 7 per cent. and 15 per cent.—1 point; responses appearing from 3 per cent. to 6 per cent.—2 points; responses appearing 1 per cent. or 2 per cent.—3 points;

responses appearing less than 1 per cent—4 points. This procedure was adopted for all the ' divergent ' tests except Hidden Figures, in which case the number of correct responses was recorded, and Symbol Production and Associations, in which any appropriate response was accepted.

Procedure.

The thirteen tests were administered to the 320 subjects on the same day, and Pearson Product-Moment correlations calculated between the thirteen sets of scores. The matrix of intercorrelations was then subjected to princial axis factor analysis, with unities in the diagonals, according to Householder's (1938) method. Rotations were carried out on the first two factors, first according to the Varimax (Kaiser, 1958) criterion for simple structure, and subsequently according to the Procrustes method (Hurley and Cattell, 1962), in order to obtain the best possible approximation to a factor completely defined by the ' divergent ' tests. The Procrustes technique involves rotating the factor matrix to the best possible fit to an hypothesized factor structure, and does not necessarily produce orthogonal factors. In the present case, the hypothetical factor structure to which the best possible fit was sought involved a first factor deriving all its variance from the ' convergent ' tests, and a second factor completely defined by ' divergent ' tests.

III.—RESULTS.

The matrix of intercorrelations among the thirteen tests is shown in Table 1.

TABLE 1

CORRELATION MATRIX.

Tests	1	2	3	4	5	6	7	8	9	10	11	12	13
Verbal I.Q.	1·00	608	671	399	771	197	362	194	351	268	145	263	394
Performance I.Q.		1·00	472	292	615	222	303	231	350	302	171	364	281
Vocabulary			1·00	409	651	132	209	135	263	302	068	228	359
Inferences				1·00	442	252	029	038	199	156	105	138	191
Academic Average					1·00	250	402	163	430	345	185	299	398
Length Estimation						1·00	114	068	128	185	139	158	055
Seeing Problems							1·00	162	408	273	188	105	170
Tin Can Uses								1·00	257	168	220	065	069
Consequences									1·00	356	297	156	200
Symbol Production										1·00	258	195	132
Circles											1·00	012	148
Hidden Figures												1·00	208
Associations													1·00

All decimal points have been omitted except in the diagonal.

When this matrix was factorized, five factors with latent roots greater than 1 were obtained. However, of these, only the first (latent root$=4{\cdot}397$), and the second (latent root$=1{\cdot}409$) were clearly related to convergent or divergent thinking, and subsequent factors were, therefore, disregarded.

Table 2 shows the two largest factors rotated according to the Varimax and Procrustes criteria.

TABLE 2

ROTATED FACTOR MATRIX.

Test	Orthogonal Factors			Oblique Factors	
	I	II	h^2	I	II
Verbal I.Q.	823	249	739	826	061
Performance I.Q.	675	333	570	632	196
Vocabulary	797	100	645	851	—101
Inferences	618	—034	384	699	—205
Academic Average	825	318	781	804	140
Length Estimation	285	187	116	251	136
Seeing Problems	220	599	407	032	622
Tin Can Uses	020	567	321	—179	639
Consequences	269	683	539	056	706
Symbol Production	256	561	380	085	570
Circles	—025	659	434	—260	759
Hidden Figures	455	054	210	487	—062
Associations	523	098	283	547	—030
Sums of Squares	3·58	2·23	5·81	3·65	2·33
Percentage Total Variance	27·5	17·1	44·6		
Percentage Common Variance	61·6	38·4	100		

All decimal points have been omitted in factor loadings.

In the case of the orthogonally rotated factor matrix, the first factor was defined by the scores for ' divergent ' processes. It was characterized by high loadings of Academic Average, Verbal I.Q., Vocabulary, Performance I.Q. and Inferences, and derived 81 per cent. of its variance from the two I.Q. measures plus Academic Average. The second factor, in the orthogonal matrix was defined by the scores for Seeing Problems, Tin Can Uses, Circles, Consequences, and Symbol Production, and derived 85 per cent. of its variance from these five ' divergent ' tests. The sixth and seventh ' divergent ' tests (Hidden Figures and Associations) had almost zero correlations with this factor and were disregarded in the further interpretation of the data.

Although the second factor resulting from the orthogonal rotation predominantly represented the ' divergent ' tests, it was by no means uncontaminated with significant loadings derived from the ' convergent ' measures. In fact, 12·3 per cent. of its variance was derived from Verbal I.Q., Performance I.Q., and Academic Average. Thus, orthogonal rotations suggested that, although ' divergent ' tests do cluster to some extent, the extent of this clustering is far less complete than some previous writers have assumed.

Consequently, an attempt was made by the use of oblique rotations, to obtain a ' pure ' ' divergent ' factor, as nearly as possible completely defined by the five most promising ' divergent ' tests, while independent of ' convergent ' tests and school achievement. The oblique loadings thus obtained are shown in Table 2. The first oblique factor loaded highly on Vocabulary, Verbal I.Q., Academic Average, Inferences, and Performance I.Q., and clearly was again a ' convergent' factor, while the second oblique factor, defined by the high loadings of Circles, Consequences, Tin Can Uses, Seeing Problems, and Symbol Production, was again a ' divergent factor.' The effect of this oblique rotation was to reduce the factor loadings of tests involving convergent processes ; nevertheless, the loadings were only reduced, not eliminated, so that the second factor was still partly dependent on convergent processes.

Oblique factors are, of course, correlated, and the correlation between the ' convergent ' factor and the ' divergent ' factor was calculated and found to be ·514. Hence, although a factor chiefly dependent on divergent process was obtained, it was significantly correlated with the factor defined by the more conventional ' convergent ' processes.

IV.—DISCUSSION.

A finding which should be mentioned in passing is connected with the failure of the R.A.T.-type test (Mednick, 1962) to correlate highly with either the orthogonal or the oblique ' divergent ' factors. Despite Mednick's claim that the R.A.T. is an associatively-based test of creativity, it correlated significantly with Academic Average (·398), and Verbal I.Q. (·394), and had near zero loadings on the ' divergent ' factors. This finding is consistent with a finding of Ketcham and Kheiralla (1962) who reported that the R.A.T. correlated significantly with measures of achievement, and casts some doubt on the usefulness of the R.A.T. as a creativity measure.

As Pribram (1964) has pointed out, there is no neurological basis for the notion of two kinds of intellectual functioning. Although the latter point of view has been proposed by some authors (Guilford, 1950 ; Getzels and Jackson, 1962 ; Torrance, 1962), the present findings do not support them. When orthogonal rotations were carried out, a factor which loaded highly on the ' divergent ' tests was obtained but, as was the case with Thorndike's analysis of the Getzels and Jackson data, this factor was still partly dependent on conventional ' convergent ' measures for its variance.

Thus, the present findings suggest that it would be wrong to argue either that convergent and divergent thinking cannot be distinguished from each other factorially, as some authors have suggested, or that they are completely independent of each other as has been suggested by others. The position seems to be similar to that suggested by Thorndike (1963b) ; it is possible to compile a set of effective tests specifically to measure divergent thinking, but the factor underlying these tests bears an oblique, not an orthogonal relation to the factor underlying intelligence tests of a conventional type. This point of view is further supported by the presence, before rotation, of a large general factor, which accounted for 76 per cent. of the common variance. It is unacceptable to think of creativity as a separate basic intellective mode.

Behaviourally, this finding indicates that high scores on tests of divergent thinking will tend to be accompanied by high scores on tests of convergent thinking, and *vice versa*, although the two kinds of thinking are not identical, and it supports the suggestion that conventional skills may provide the basis upon which creative productions rest, as Bruner (1962) and Pribram (1964) have suggested.

V.—REFERENCES.

ANDERSON, C. C. (1964). *Originality, Impulsive Extraversion and Metaphor.* Unpublished paper, University of Alberta.

ANDREWS, E. G. (1930). The development of imagination in the pre-school child. *University of Iowa Studies in Character*, 3.

BARRON, F. (1955). The disposition toward originality. *J. Abnorm. Soc. Psychol.*, 51, 478-485.

BARRON, F. (1963). *Creativity and Psychological Health.* New York : Von Nostrand.

BRUNER, J. S. (1958). The conditions of creativity. In Gruber, H. E., Terrell, G., and Wertheimer, M. (Eds.) (1962) *Contemporary Approaches to Creative Thinking.* New York: Atherton Press.

BURT, C. (1962). Creativity and intelligence (Critical Notice). *Brit. J. Educ. Psychol.*, 32, 292-298.

BURT, C. (1964). Foreword. In Koestler, A., *The Act of Creation.* London · Hutchinson.

COUCH, A., and KENISTON, K. (1960). Yeasayers and naysayers: agreeing response set as a personality variable. *J. Abnorm. Soc. Psychol.*, 60, 151-174.

CROPLEY, A. J. (1965*a*). *Originality, Intelligence and Personality.* Unpublished doctoral dissertation, University of Alberta.

CROPLEY, A. J. (1965*b*). The relatedness of convergent and divergent thinking. *Alta. J. Educ. Res.*, 11, 176-181.

DREVDAHL, J. E., and CATTELL, R. B. (1963). *Personality and Creativity in Artists and Writers.* Mimeo paper.

FRENCH, J. W, Ekstrom, R. B, and PRICE, L. A. (1963). *Manual for Kit of Reference Tests for Cognitive Factors.* Princeton: E.T.S.

FREUD, S. (1947). *Leonardo da Vinci: A Study in Psychosexuality* (1910). New York: Random House.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development.* London: Dent.

GETZELS, J. W., and JACKSON, P. W (1962). *Creativity and Intelligence.* New York: Wiley.

GHISELIN, B. Ultimate criteria for two levels of creativity. In Taylor, C. W. (Ed.) (1959). *The Second* (1957) *University of Utah Research Conference on the Identification of Creative Scientific Talent.* Salt Lake City: University of Utah Press.

GORDON, W. J. (1961). *Synectics.* New York: Harper Bros.

GUILFORD, J. P. (1950). Creativity. *Amer. Psychologist*, 5, 444-454.

GUILFORD, J. P. (1959). Three faces of intellect. *Amer. Psychologist*, 14, 469-479.

GUILFORD, J. P., and MERRIFIELD, P. R. (1960). *The Structure of Intellect Model: Its Uses and Implications.* (Report of the Psychological Laboratory, No. 24). Los Angeles: University of Southern California Press.

HOUSEHOLDER, A. S., and YOUNG, G. (1938). Matrix approximation and latent roots. *Ann. Math. Stat.*, 45, 165-171.

HURLEY, J. R., and CATTELL, R. B. (1962). The Procrustes program: producing direct rotation to test a hypothesized factor structure. *Behav. Sci.*, 7, 258-262.

KAISER, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.

KETCHAM, W. A., and KHEIRALLA, S. (1962). Creativity in relation to intelligence and school achievement. In *Collected Papers of the Inter-Institutional Seminar.* Education Department, Henry Ford Museum.

KOESTLER, A. (1964). *The Act of Creation.* London: Hutchinson.

LOWENFELD, V. (1958). Current research on creativity. *J. Nat. Educ. Assoc.*, 47, 538-540.

McGUIRE, C., HINDSMAN, E., KING, F. J., and JENNINGS, E. (1961). Dimensions of talented behaviour. *Educ. Psychol. Measmt.*, 21, 3-38.

McNEMAR, Q. (1964). Lost: our intelligence? Why? *Amer. Psychologist*, 19, 871-882.

MACKLER, B. (1962). *Creativity and Life Styles.* Unpublished doctoral dissertation, University of Kansas.

MALTZMANN, I. (1960). On the training of originality. *Psychol. Rev.*, 67, 229-242.

MARSH, R. W. (1964). A statistical reanalysis of Getzels' and Jackson's data. *Brit. J. Educ. Psychol.*, 34, 91-93.

MEDNICK, S. A. (1962). The associative basis of creativity. *Psychol. Rev.*, 69, 220-232.

MEER, B., and STEIN, M. I. (1955). Measures of intelligence and creativity *J. Psychol.*, 39, 117-126.

PRIBRAM, K. H. (1964). Neurological notes on the art of educating. In Hilgard, E. R. (Ed.) *N.S.S.E. Yearbook*, LXIII, Chicago: University of Chicago Press.

RICHARDS, J. M., CLINE, V. B., and NEEDHAM, W. E. (1964). Creativity tests and teacher and self-judgments of originality. *J. Exp. Educ.*, 32, 281-285.

SULTAN, E. E. (1962). A factorial study in the domain of creative thinking. *Brit. J. Educ. Psychol.*, 32, 78-82.

TAYLOR, C. W. (Ed.) (1964). *Creativity: Progress and Potential.* New York: McGraw-Hill.

TERMAN, L. M. (1925). *Genetic Studies of Genius.* Vol. I, Stanford: Stanford University Press.

Thorndike, R. L. Some methodological issues in the study of creativity. In Gardner, E. F. (Ed.) *Proceedings of the* 1962 *Invitational Conference on Testing Problems.* Princeton : E.T.S., 1963a.

Thorndike, R. L. (1963b, 64). The measurement of creativity. *Teach. Coll. Rec.*, 422-424.

Torrance, E. P. (1962). *Guiding Creative Talent.* Englewood Cliffs : Prentice-Hall.

Torrance, E. P. (1963). *Education and the Creative Potential.* Minneapolis : University of Minnesota Press.

Toynbee, A. J. (1962). Has America neglected its creative minority ? *Calif. Monogr.*, 72, 7-10.

Welsh, L. (1946). Recombination of ideas in creative thinking. *J. Appl. Psychol.*, 30, 638-643.

Wodtke, K. H. (1964). Some data on the reliability and validity of creativity tests at the elementary school level. *Educ. Psychol. Measmt.*, 24, 399-408.

Yamamoto, K. (1965). Effects of restriction of range and test unreliability on correlation between measures of intelligence and creative thinking. *Brit. J. Educ. Psychol.* 35, 300-305.

(*Manuscript received* 18*th October*, 1965)

PUPIL-TEACHER RELATIONSHIPS IN CO-EDUCATIONAL AND SINGLE-SEX GRAMMAR SCHOOLS

By R. R. DALE

(*Senior Lecturer in Education, University College of Swansea*)

SUMMARY. The sample population was some 2,000 ex-grammar school pupils who were intending teachers ; it was, therefore, representative of only a section of the grammar school population. A previous article based on 1,500 of this sample reported that the co-educated women found the school atmosphere pleasanter and life in school happier than did the women educated in girls' schools ; the co-educated men had similar but slighter trends, significant only for school atmosphere. The co-educated men and women expressed a very much greater preference for co-education than their opposite numbers did for single-sex education. The present article describes a first attempt to discover the reasons for these opinions, in the region of pupil-teacher relationships. The co-educated women thought their teachers more friendly, more helpful, more influential for good and less influential for bad, than did the women from girls' schools. These results support and extend the previous findings.

I.—INTRODUCTION.

IN this *Journal* (36, February, 1966), the writer described an enquiry into the comparative happiness, as pupils at mixed and single-sex grammar schools, of a group of 1,500 students who were then in departments of education and training colleges. The co-educated women reported a higher level of happiness at school and estimated the school atmosphere to be pleasanter than did those women who had attended girls' schools. The corresponding differences between the two groups of men were in the same direction, but appreciably smaller, only their results on atmosphere being statistically significant.* Support for both the general direction of the findings, in favour of co-education, and for the tendency for the women's groups (co-educated and single-sex educated) to differ more than the men's groups, was given by the results from a third question. When the co-educated groups were asked whether they would have preferred a single-sex schooling, and the other groups were asked the reverse question, 96 per cent. of the co-educated women chose co-education, and 53 per cent. of the women from girls' schools (with 22 per cent. undecided), while the difference between the men's groups was smaller, though still large, 91 per cent. of the co-educated men preferring co-education but only 42 per cent. of the men from boys' schools (with 19 per cent. undecided). The present article gives the results of further questions intended to discover some of the reasons underlying the preference.†

II.—PROCEDURE.

A questionnaire was administered to some 2,000 intending teachers who were in training colleges and departments of education, and who were ex-pupils of grammar schools. The replies from women and men were analysed separately. The proportion of students from departments of education included in the sample was roughly the same as that in the ' population ' of intending teachers. The questionnaire was anonymous and was administered to large groups at a time under examination conditions ; the percentage of replies was, therefore, 100 per cent. It was exploratory in nature, designed to cover a broad area of attitudes which had not been previously investigated, to enable future research workers to concentrate on the most interesting of the results. Four of the questions were designed to assess the general attitude of the ex-pupils towards their teachers. The first asked " Do you consider that your teachers, judged as a whole, were friendly ? " Respondents were asked to indicate their value-judgment on a five-point scale ranging from ' Very friendly,' through ' Neutral ' to ' Very unfriendly.' The second was " Do you consider that your teachers, judged as a whole, were helpful, apart from class teaching ? " Similar value-judgments were requested. The third was " Do you consider that any one (or several) of your teachers had a powerful influence for good in the development of your personality ? " Here the scale was ' Decidedly,' ' Possibly,' and ' No.' The fourth question resembled the third but substituted the word ' bad ' for ' good.' A space for comment was left after the last two questions, both to uncover ambiguities or other faults and in order to ascertain the more important reasons which impelled the respondents to answer ' Yes.' It would have been desirable for a similar opportunity for comment to have been left after the other two questions, but limitation of time prevented this.

---

* The later addition of 500 extra respondents brought the figure up to the significance level (P= ·025).

† Cf. also Dale (1955 and 1965), and Davies, W. E. (1950).

III.—RESULTS.

Friendliness of Teachers.

The replies of the men to the question on the friendliness of teachers are included in Table 1. Though the co-educated men thought their teachers were slightly more friendly than did those educated in boys' schools, the difference is not statistically significant. In both groups only some 3 to 4 per cent. of ex-pupils thought their teachers ' Rather unfriendly ' or worse. On the other hand, rather more than a fifth of both groups answered ' Neutral.'

TABLE 1

FRIENDLINESS OF TEACHERS.

	Replies from Men ex-pupils of :				Replies from Women ex-pupils of :			
	Boys' schools		Co-educational schools		Girls' schools		Co-educational schools	
	N	%	N	%	N	%	N	%
Very friendly	84	14·5	63	18·1	52	8·2	94	23·2
Friendly	345	59·6	199	57·2	379	59·4	248	61·2
Neutral	129	22·3	76	21·8	181	28·4	59	14·3
Rather unfriendly	19	3·3	9	2·6	22	3·4	5	1·2
Very unfriendly........	2	0·3	1	0·3	4	0·6	0	0
Totals	579	100%	348	100%	638	100%	406	100%

In the case of the women the difference between the co-educated group and that educated in girls' schools is much more marked than that between the men's groups, and statistically significant, more than a fifth of the co-educated women assessing their teachers as 'Very friendly,' and only a twelfth of the women from girls' schools. In neither group of women is there more than 4 per cent. who thought their teachers 'Rather unfriendly' or worse, though well over one-quarter of the ex-pupils of girls' schools and over one-eighth of the co-educated placed their teachers in the 'Neutral' category. Perhaps the most interesting contrast comes when we combine the 'Very friendly' and 'Friendly' categories: this embraces 68 per cent. of respondents from girls' schools and 84 per cent. of those from co-educational schools. The difference between the two entire distributions is statistically highly significant.

Helpfulness of Teachers.

The question on the helpfulness of teachers produced the same pattern of results (Table 2) as that derived from the previous question. The difference between the co-educated men and those from boys' schools favours the co-educational schools but is small and statistically not significant.* Rather more ex-pupils (some 8 per cent. of the co-educated and 6 per cent. from boys' schools) thought their teachers 'Rather unhelpful' or worse, but this is a small proportion. In both cases rather more than a fifth gave the answer 'Doubtful,' while the great majority—70 per cent.—recorded 'Very helpful' or 'Helpful.'

TABLE 2

HELPFULNESS OF TEACHERS.

	Replies from Men ex-pupils of:				Replies from Women ex-pupils of:			
	Boys' schools		Co-educational schools		Girls' schools		Co-educational schools	
	N	%	N	%	N	%	N	%
Very helpful	130	22·5	94	27·2	100	15·6	114	28·1
Helpful	278	48·0	149	43·2	297	46·4	215	53·0
Doubtful	135	23·3	73	21·2	183	28·5	56	13·8
Rather unhelpful	28	4·8	26	7·5	55	8·6	16	3·9
Very unhelpful	9	1·6	3	0·9	6	0·9	5	1·3
Totals	580	100%	345	100%	641	100%	406	100%

The results from the two groups of women resemble closely the results from the previous question. Well over one-quarter of the co-educated women assessed their teachers as 'Very helpful,' but less than one sixth of the ex-pupils of girls' schools. If we combine the 'Very helpful' with the 'Helpful, 'while both groups have a high percentage, the co-educational group reaches 81 per cent., while the other group has 62 per cent. Well over one-quarter of the ex-pupils of girls' schools recorded 'Doubtful,' but the percentage for the co-educated women was not quite a seventh. Though only a small proportion of either group classed their teachers in the two unhelpful categories, there was almost twice the percentage recorded by the women from girls' schools as by those from co-educated schools. Results from the chi-square test show that it is highly unlikely that such a difference in the distributions would have occurred by chance.

---

* Chi-square=4·625 for 3 degrees of freedom.

Influence for Good or Bad.

In their replies about whether one or more teacher had had a powerful influence for good or for bad, the co-educated and single-sex educated groups, of both sexes, were much closer together. In the case of the men, the proportions of the two groups recording 'Decidedly,' 'Possibly,' and 'No,' for 'good influence' were almost identical. There was likewise little divergence between them under 'distinctly bad influence.' About a tenth of both groups considered that one or more teachers had had a 'distinctly bad influence'; on the other hand, over one-third gave the same answer for 'Powerful influence for good.' The women's groups were somewhat further apart and the differences were statistically significant.

The comments made by respondents revealed no difficulties or ambiguities in the questions. When the classification of them is completed, it is hoped that they will be the subject of a further article.

TABLE 3A

INFLUENCE OF ONE OR MORE TEACHERS: REPLIES FROM MEN.

	Powerful influence for good				Powerful influence for bad			
	Boys' schools ex-pupils		Co-educational schools ex-pupils		Boys' schools ex-pupils		Co-educational schools ex-pupils	
	N	%	N	%	N	%	N	%
Decidedly	213	37·2	119	34·5	59	10·3	34	9·8
Possibly	252	44·0	165	47·8	103	18·0	66	19·1
No	108	18·8	61	17·7	409	71·7	246	71·1
Totals	573	100%	345	100%	571	100%	346	100%

TABLE 3B

INFLUENCE OF ONE OR MORE TEACHERS: REPLIES FROM WOMEN.

	Powerful influence for good				Powerful influence for bad			
	Girls' schools ex-pupils		Co-educational schools ex-pupils		Girls' schools ex-pupils		Co-educational schools ex-pupils	
	N	%	N	%	N	%	N	%
Decidedly	148	23·4	121	29·9	62	9·9	23	5·7
Possibly	310	49·1	205	50·6	95	15·2	49	12·2
No	174	27·5	79	19·5	469	74·9	329	82·1
Totals	632	100%	405	100%	626	100%	401	100%

IV.—DISCUSSION.

The results from all four questions show the same trends and help to confirm each other. In all of them, the differences between the two groups of women are greater than those between the groups of men, and all are in the same direction. Exactly the same trends have been reported in previous work by the writer and are referred to in the Introduction.

We now need to ask why the co-educated women recorded their teachers as more friendly, more helpful, with a greater influence for good and less influence ' for bad,' than did those women who were educated in girls' schools. It may be, of course; that this gives a true picture of the situation. On the other hand, the argument could be put forward that the co-educated women either consciously or unconsciously preferred being taught by men ; this favourable emotional outlook might, therefore, lead them to give a higher estimation of their teachers' friendliness, etc., even though the teachers were, in fact, no more friendly than those in a girls' school. On the other hand, even if this be so, the pupils themselves would presumably be affected beneficially even by thinking that their teachers were more friendly. Another possibility, though perhaps remote, is that the type of education offered in a girls' school leads to a more restrained judgment than does the education of a co-educational school. This last hypothesis could be countered, at least in part, by another research finding, namely that an overwhelming majority of those teachers who have experienced co-education express a strong preference for it (Dale, 1955) ; this being so, they are, presumably, happier than their counterparts in single-sex schools and, therefore, would feel inclined to be more helpful and more friendly towards their pupils.

The differences between the two men's groups, although in the same direction, are generally small and lack statistical significance ; the largest of them results from the question on the helpfulness of the teachers. Here the same arguments will apply to the men as were used above about the women. Other relevant considerations are that a greater proportion of the co-educational grammar schools have semi-rural settings, and the more attractive surroundings might make pupils and teachers happier, thereby increasing their estimates on such qualities as ' Friendliness of staff.' Against this could be set the argument that only a tiny fraction of the co-educational schools are ' Direct Grant ' and similar high status schools, and this factor should increase the happiness of both the pupils and the staffs of these single-sex schools. Finally, readers should again be reminded that the sample is one of intending teachers who are ex-pupils of grammar schools. They will, therefore, be of higher academic quality than the average, and probably more inclined to enjoy their school days and record high estimates of their teachers.

V.—REFERENCES.

DALE, R. R. (1955). Co-education : the verdict of experience *Brit. J. Educ. Psychol.* 25, 3.

DALE, R. R. (1965). Co-education : the verdict of experience, II—A qualitative approach *Brit. J. Educ. Psychol.*

DALE, R. R. (1966). The Happiness of Pupils in Co-educational and Single Sex Grammar Schools, *Brit. J. Educ. Psychol.*, 36, 1.

DAVIES, W. E. (1950). *A Study of the Attitudes of Secondary School Teachers Towards Co-education, Differentiation of Curriculum and Sex Teaching* Ph.D. Thesis, London.

(*Manuscript received 7th October, 1965*)

SOCIAL AND EDUCATIONAL VARIABLES RELATING TO TEACHERS' ASSESSMENTS OF PRIMARY SCHOOL PUPILS

By D. McINTYRE, A. MORRISON and J. SUTHERLAND
(*Moray House College of Education, Edinburgh*)

SUMMARY. Teachers' ratings of the personality traits and attainments of their pupils were analysed according to the social class background of schools, age and experience of teachers, and teachers' opinions about education Examination of the results of the principal components analyses and Varimax rotations showed that assessments were made on two main dimensions, but that the traits defining these dimensions differed according to the social class of the pupils, their sex, and the age and experience of the teachers. Assessments for girls were largely uniform across the different analyses, the first factor emphasising traits descriptive of good behaviour. For boys, however, there were considerable social class differences, both in the relative importance of factors and in the traits descriptive of the kind of pupil most worth taking trouble over. Teachers in middle class and mixed social class schools laid stress on traits descriptive of the pleasant and trustworthy pupil whilst those in urban and suburban working class schools showed particular concern for the pupils' attainments and attitudes to school. Assessments by older and more experienced teachers emphasised attainment and attitudes to work whilst those of younger teachers were primarily concerned with the good behaviour of pupils.

I.—INTRODUCTION.

OBSERVATIONS of teachers' attitudes to their work and to their pupils suggest the possible importance of the social class of pupils and of the teachers' opinions about education as variables affecting ways in which pupils are assessed. Craig (1960), in a comparative study of teachers in Scottish junior and senior secondary schools, found that there were considerable differences between the two groups in their opinions of their pupils and their views on the teachers' functions. Also, previous studies of teachers' personality ratings of pupils in secondary and primary schools (e.g., Hallworth, 1961, and Morrison, McIntyre and Sutherland, 1965) have shown variations in assessments associated with types of schools, age and sex of pupils, and sex of teachers. This article reports on the second stage of a study of teachers' assessments of pupils in a number of Scottish primary schools, examining personality trait and attainment assessments made in schools of differing social class background, and by teachers in different age groups, and with contrasting opinions about education.

II.—DESIGN OF INVESTIGATION.

Since the experimental design and statistical treatment have been fully described in the previous article (Morrison, et *al.*, 1965) only those aspects which are particularly relevant to this stage of the study are mentioned here. Thirty-four teachers of 11–12-year-old pupils in a Scottish city were chosen so as to give a balanced representation of the following variables :

(*a*) Social class of school, in four categories :
Middle class (M.C.).
Mixed social class (M.S.C.).
Suburban working class (S.W.C.).
Urban working class (U.W.C.).

(*b*) Age, in three categories: under 40 years of age, 40–50 years of age, and over 50 years of age.

(*c*) Sex/marital status, in three categories: men, married women and single women.

Each teacher rated boys and girls independently on twenty-five traits. Three further measures were also included for analysis: Sociometric Status, I.Q. and School Attendance. Also, each teacher completed the revised form of the Manchester scales of Opinions about Education (Oliver and Butcher, 1962). Sixty-eight correlation matrices were calculated for the ratings obtained from each of thirty-four teachers, and these were then combined in various groups in order to examine the influence of the particular social and educational variables under discussion.

(A) Social class: all the teachers were grouped, and their matrices combined, according to the four social class categories of schools, giving sub-sample sizes of 9, 9, 8 and 8, respectively.

(B) Age: Eight teachers over the age of 55 years, and each with over 35 years of teaching experience were grouped and compared with seven teachers aged 37 or less, and each with less than fourteen years of experience.

(C) Opinions about Education: each teacher completed three scales, two of which proved useful; these being designed to measure Radicalism/Conservatism (R scale) and Tough/Tendermindedness (T scale), respectively. On the basis of the two scores obtained from each of the thirty-four teachers, between 5 and 7 (depending on the distribution of scores) were chosen to represent the extremes on each of the two scales.

III.—RESULTS.

Since the analyses carried out for the sex/marital status comparisons showed that in general only three components of any importance are obtained, it was decided that three principal components shoould be extracted. These were then subjected to Varimax-rotations.*

A.—*Social Class.*

Boys: In all cases there is one factor which is considerably larger than the others, and which defines what may be generally described as a 'good pupil,' with different aspects of behaviour predominating in each social class category. These factors are shown in abbreviated form in Table 1.

A feature of the boys' analyses not shown in the abbreviated Table 1 is that the loadings on Leadership, Popularity and Social Confidence, which are highest for the suburban working class, are also considerable for mixed social class, a good deal smaller for the middle class, and negligible for the urban working class. Loadings on I.Q. are ·68 for S.W.C., ·50 for U.W.C., ·39 for M.S.C., and ·20 for M.C., and loadings for attainment ratings of English and Arithmetic follow the same general pattern of decreasing importance.

In comparing the dimensions defined by the first factors, it is noteworthy that in all cases there are very high loadings on 'pupil worth taking trouble over.' Another striking feature is the greater importance given to attainment by the teachers in working class schools. In the middle class schools, teachers appear to consider the 'pleasantness' of a boy—the degree to which he conforms to the norms of classroom behaviour—as the attribute which should most

---

* Copies of complete results may be obtained from the writers.

TABLE 1

BOYS: MAIN LOADINGS ON VARIMAX FACTOR 1.

Middle class		Mixed social class		Suburban working class		Urban working class	
Trustworthiness	·84	Pleasantness	·85	English attainment	·88	Attentiveness	·82
Co-op. with teacher	·80	Trustworthiness	·83	Pupil worth trouble over	·82	Pupil worth trouble over	·77
Courtesy	·80	Attentiveness	·83	Arithmetic attainment	·81	Co-op. with teachers	·71
Pupil worth trouble over	·77	Courtesy	·81	Originality	·78	Enthusiasm	·71
Pleasantness	·77	Enthusiasm	·80	Enthusiasm	·77	English attainment	·69
Attentiveness	·75	Co-op. with teachers	·79	Home background	·76	Persistence	·66
Carefulness	·75	Carefulness	·76	Leadership	·74	Originality	·65
Gentleness	·74	Persistence	·75	Social confidence	·73	Trustworthiness	·60
Enthusiasm	·73	Pupil worth trouble over	·75	Persistence	·73	Arithmetic attainment	·59 ·59
Co-op. with pupils	·72	Popularity	·72	Popularity	·72	Home gack-background	·58
Persistence	·67	Home background	·68	Trust-worthiness	·69	Carefulness	·52
% Variance extracted: *	33		38		39		25

* In this and subsequent abbreviated tables, the percentages of variance extracted refer to the loadings on all the measures.

earn the teacher's interest and concern for him. The teachers in working class schools, and particularly in suburban working class schools, consider that for a boy to be specially deserving of their attention he should show high attainment and originality.

The second factor in the boys' analyses is concerned with aspects of sociability and social leadership. However, in the case of the middle class group this factor also has high loadings on English and Arithmetic Attainment and Originality. Third factors vary considerably across the analyses; the largest, for U.W.C. boys, is characterised by such traits of good behaviour as Calmness, Courtesy, Quietness and Co-operation with teachers, whilst that for S.W.C. is similar, the dominant loadings being Gentleness and Pleasantness. M.S.C. boys have a third factor of I.Q., Originality and Attainment, and the M.C. boys one of I.Q.

The analyses for girls, as for boys, show two main factors. The first, given in Table 2, is descriptive of the 'good pupil,' while the second is concerned with sociability and social leadership. However, assessments of girls differ from those of boys in that they are more uniform across analyses and generally extract more variance. Third factors are all small and tend to be dominated by one trait—Cheerfulness for M.C. girls, Sociometric status for the U.W.C. pupils, and I.Q. for S.W.C. and M.S.C. groups.

B.—*Age of Teachers.*

The first factors for boys and girls are shown in Table 3.

TABLE 2

GIRLS: MAIN LOADINGS ON VARIMAX FACTOR 1.

Middle class		Mixed social class		Suburban working class		Urban working class	
Courtesy	·86	Trustworthiness	·91	Pleasantness	·82	Co-op. with teachers	·93
Co-op. with teacher	·86	Courtesy	·89	Co-op. with teachers	·82	Attentiveness	·88
Trustworthiness	·84	Co-op. with teachers	·89	Turstworthiness	·82	Trustworthiness	·87
Attentiveness	·84	Attentiveness	·88	Pupil worth trouble over	·81	Courtesy	·87
Carefulness	·82	Pleasantness	·84	Popularity	·79	Pupil worth trouble over	·80
Enthusiasm	·77	Pupil worth trouble over	·81	Enthusiasm	·77	Quietness	·78
Pupil worth trouble over	·76	Quietness	·76	Personal appearance	·77	Persistence	·77
Persistence	·76	Carefulness	·76	Home background	·76	Careulness	·75
Gentleness	·75	Gentleness	·76	Persistence	·74	Pleasantness	·74
Co-op. with pupils	·73	Persistence	·75	Attentive	·74	Enthusiasm	·73
Home background	·71	Home background	·75	Speech	·73	Gentleness	·73
Pleasantness	·70	Personal appearance	·74	Arithmetic attainment	·70	Calmness	·70
% Variance extracted	·39		41		43		39

TABLE 3

BOYS AND GIRLS RATED BY OLD AND BY YOUNG TEACHERS.
MAIN LOADINGS ON VARIMAX FACTOR 1.

BOYS				GIRLS			
Old		Young		Old		Young	
Pupil worth trouble	·85	Co-op. with teachers	·88	Social confidence	·81	Attentiveness	·85
Persistence	·83	Courtesy	·84	Leadership	·78	Persistence	·79
English attainment	·82	Attentiveness	·80	English attainment	·77	Courtesy	·77
Carefulness	·81	Pleasantness	·77	Originality	·76	Trustworthiness	·76
Attentiveness	·76	Trustworthiness	·74	Persistence	·73	Carefulness	·72
Arithmetic attainment	·72	Pupil worth trouble	·72	Arithmetic attainment	·71	Co-op. with teacher	·69
Enthusiasm	·71	Co-op. with pupils	·67	Co-op. with other pupils	·71	Pupil worth trouble	·69
Trustworthiness	·69	Enthusiasm	·63	Pupil worth trouble over	·70	Quietness	·66
Originality	·67	Persistence	·62	Games ability	·69	English attainment	·62
Home background	·64	Home background	·61	Popularity	·66	Speech	·62
Personal appearance	·64	Speech	·60	Home background	·64	Calmness	·60
I.Q.	·63	Calmness	·60	Enthusiasm	·64	Enthusiasm	·59
				Personal appearance	·64	Personal appearance	·59
				Speech	·63	Arithmetic attainment	·58
				I.Q.	·61	Gentleness	·58
						Pleasantness	·57
% Variance	34		30		36		29

While the first factors for young teachers' ratings of boys and of girls are very similar, detailed comparison of all four sets of results is made easier by considering these factors together with the second factors.

For both analyses of girls there are large second factors. The second factor for old teachers (24 per cent. variance) is comparable with the first factor for young teachers, but has lower loadings on such work-directed traits as Carefulness, Attentiveness, Persistence, Enthusiasm, English Attainment and Speech. It does, however, have a higher loading on I.Q. Similarly, the second factor for young teachers (27 per cent. variance) can be compared with the first factor for old teachers. It is similar in having high loadings on social leadership traits, but places more emphasis on Popularity and less on English attainment and Persistence. It also has much lower loadings on Attentiveness, Trustworthiness and Co-operation with Teachers. The third factors for both analyses of girls are small and in both cases have high loadings on Gentleness and Quietness.

The second factor for the young teachers' ratings of boys (18 per cent. variance) has highest loadings on English and Arithmetic Attainment, Originality, I.Q., Enthusiasm, Carefulness and Home Background, with a load;ng of ·46 on 'Pupil worth taking trouble over.' The third factor in this analysis opposes high loadings on Popularity, Leadership, Social Confidence and Cheerfulness to those on Quietness and Gentleness. For the old teachers' ratings of boys, the two opposing ends of the young teachers' sociability and social leadership factor separate into two independent factors, one defined by Cheerfulness, Popularity, Pleasantness, Co-operation with Other Pupils, Leadership, Games Ability and Social Confidence, and accounting for 15 per cent. of the variance; the other (11 per cent. variance) being defined by Quietness, Gentleness and Carefulness, with a relatively small opposed loading on Games Ability.

C.—*Opinions About Education.*

Variables associated with teachers' opinions.—Whilst the experiment was designed to ensure the mutual independence of the variables of social class, age and sex/marital status, it remained necessary to discover what relationships, if any, existed between these variables and the scores of teachers on the two opinion scales for radicalism/conservatism and tough/tendermindedness. Analyses of variance were, therefore, carried out. A significant relationship was found between social class and T-scale score ($p < \cdot 05$), teachers in S.W.C. schools being most tenderminded and those in U.W.C. schools being most tough-minded. Also both scales were significantly related to age, older teachers being more toughminded ($p < \cdot 1$) and more conservative ($p < \cdot 05$). These relationships between social class and T scale score, and between age and scores on both the R and T scales, were apparent in the composition of the four sub-samples of raters with extreme scores on each of the two scales.

Tough-tendermindedness. —The analyses of the two groups of teachers for boys and girls show a high degree of uniformity with regard to the patterns of loadings. For both groups of teachers rating boys there is a large factor, accounting for 36 per cent. of the variance in both cases, of high attainment and good behaviour, a second factor of sociability and social leadership (19 per cent. and 17 per cent. of the variance), and a negligible third factor. The only differences are on the first factor where toughminded teachers place strong emphasis on Quietness as an aspect of good behaviour, while tenderminded teachers treat it as irrelevant, but give rather more importance than the other group to Personal Appearance, Speech, Leadership, Physical Maturity and Popularity.

For the two analyses of girls, the toughminded group of teachers assess pupils much more in terms of the two major factors than do the tenderminded group, extracting 41 per cent. and 25 per cent. of the variance on the first two factors, respectively, as opposed to 31 per cent. and 22 per cent. for the tenderminded group. In both girls' analyses the first factor defines the well-behaved and hard-working pupil, and the second the social leader and high attainer. On the first factors there are higher loadings for the toughminded group on Quietness, Gentleness and English Attainment, while on the second factors this group has higher loadings on Physical Maturity, Leadership and Social Confidence, but a lower loading on I.Q. Only the tenderminded group has a well defined third factor, opposing Quietness and Gentleness to Leadership, Popularity, Games Ability and Social Confidence.

TABLE 4

Boys and Girls Rated by Radical and by Conservative Teachers: Main Loadings on Variance Factor 1.

Boys				Girls			
Radicalism		Conservatism		Radicalism		Conservatism	
Attentiveness	·92	Cheerfulness	·86	Persistence	·80	Leadership	·86
Pupil worth trouble over	·85	Co-op. other pupils	·85	Attentiveness	·78	Popularity	·84
Co-op. with teachers	·84	Leadership	·81	Trustworthiness	·72	Co-op. other pupils	·84
Trustworthiness	·79	Social confidence	·80	Enthusiasm	·71	Orginality	·83
Persistence	·78	Popularity	·80	Pupil worth trouble over	·68	Social confidence	·82
English attainment	·76	Orginality	·76	Carefulness	·66	Pupil worth trouble over	·78
Home background	·76	Pupil worth trouble over	·73	English attainment	·66	Persistence	
Courtesy	·75	Persistence	·70	Arithmetic attainment	·65	Enthusiasm	·76
Carefulness	·74	Enthusiasm	·70	Co-op. with teachers	·64	Arithmetic attainment	·74
Enthusiasm	·74	Pleasantness	·69	Courtesy	·64	Cheerfulness	·74
I.Q.	·69	Games ability	·63	Quietness	·60	Pleasantness	·69
Arithmetic attainment	·66	Co-op. with teachers	·59	Home background	·60	English attainment	·69
		Trustworthiness	·59	Originality	·60	Games ability	·67
						I.Q.	·67
% Variance	35		34		27		38

Radicalism-Conservatism.—As shown in Table 4, the largest factor in the analyses of radical teachers defines a well-behaved, hard-working pupil of high attainment, while those of conservative teachers define a sociable pupil with qualities of leadership, who is also hard-working and of high attainment. Correspondingly, second factors for the analyses of radical teachers have high loadings on sociability and social leadership traits, and those for conservative teachers have high loadings on traits of good behaviour. The only third factor which is either sizeable or well defined is that for radical teachers' ratings of girls, where Quietness and Gentleness are opposed to Originality, Leadership and other extravert traits.

IV.—DISCUSSION AND CONCLUSIONS.

From research on practising and student-teachers (Butcher, 1965) there is some evidence which suggests that experienced teachers tend to be more conservative and toughminded than those starting their careers ; and, in the present research, older teachers tended to be more conservative and toughminded. Thus, any conclusions concerning the effects of age and experience or of educational opinions on teachers' assessments are likely to be related and are, therefore, best considered together.

It might be hypothesised that older teachers would be more assured about handling classes, and also, because of traditional educational practices, place greater emphasis on high attainment. This seems to be borne out. The first dimensions of assessment of boys and girls used by the younger teachers stress good behaviour, while those for the older teachers are heavily loaded with attainment and such associated ' work ' traits as Persistence and Enthusiasm. Furthermore, the rating of ' Pupil worth taking trouble over ' is closely associated with high attainment only by the older teachers. Thus, despite a sex variation in the specific factor structures, several consistent differences appear between the modes of assessment used by the old and young teachers. The differences between the dimensions of assessment found for conservative and radical teachers closely parallel those between old and young with respect to the importance placed upon good behaviour, but not with respect to attainment. The interaction which has been established between age and experience and teachers' opinions makes it impossible to decide whether differences in modes of assessment should be accounted for primarily by the effects ofage and experience, with the related explanation suggested above, or by variations in radicalism-conservatism.

The T-scale is also associated with age and experience although for our sample of teachers not at a high level ($p < \cdot 1$) of significance. The differences between the old and young teachers are not, however, paralleled by a similar difference between the toughminded and tenderminded teachers. Indeed, the only considerable difference between assessments by these groups is that, as one might expect, the toughminded place much more importance on Quietness as an aspect of good behaviour.

The two most interesting features of the social class analyses for boys are concerned with attainment and with the characteristics of the ' Pupil worth taking trouble over.' Broadly, for the teachers in middle class and mixed social class schools, he is the pleasant and trustworthy boy ; for suburban working class schools, the clever boy ; and for urban working class, the boy who is attentive and hardworking. Teachers in working class schools may feel that the most effective influence they may have on pupils is to concentrate attention on those who are high attainers and/or hard workers to achieve maximum academic success, and thus to give them the opportunity to move up the social scale. On the other hand, teachers in middle class schools would seem to give most attention to those whom they find personally most pleasant.

The major difference between assessments of boys in U.W.C. and S.W.C. environments was that the predominant concern of the former was with such traits as Attentiveness and Co-operativeness, while the latter emphasised attainment which they associated with socially extravert traits. These two groups of teachers also differed considerably in their mean scores on the T-scale. Both these differences are consistent with the view that U.W.C., but not S.W.C., teachers feel they are working against an environment which is unsympathetic to the school.

More detailed consideration of all the girls' analyses shows, in the first place, that the three factors, and particularly the first two, consistently extract a greater amount of variance in the analyses for girls than do those for boys, the greatest differences being for the urban working class, toughminded and old teachers, respectively. Secondly, all the largest dimensions used for assessing girls, except for those of the old and conservative teachers, are dominated by traits of good behaviour. Finally, the greater uniformity in assessments of girls found over the different social class groups is also evident in comparing the old and young teachers, though not for either of the opinion scale comparisons.

It was reported in the previous article that " girls are perceived in a more stereotyped and in a less analytic way than boys." The present results confirm this statement. Teachers appear to make a more complete assessment in terms of one or two major dimensions of girls than they do of boys ; and they vary much less in the qualities which they look for in girls. In particular, the ' ideal girl ' is much the same whatever the social class background, while the ' ideal boy ' appears to vary according to the apparent possibilities and limitations of the environment.

In conclusion, while there is a high level of consistency in the clusters of traits which teachers group together, several variations have been found, in accordance with the social and educational variables studied, both in the ways clusters are associated and the relative importance given to them. In particular, the age and experience of the teachers, the social class background of schools and the sex of the pupils appear to interact to affect the teacher's conception of his educational role, his responses to individual pupils, and the meaning of any estimates of attainment which he may make.

V.—References.

Butcher, H. J. (1965). The attitudes of student teachers to education. *Brit. J. Soc. Clin. Psychol.*, 4, 17-24.

Craig, H. (1960). The teacher's function : some observations on an aspect of the teacher's job in Scotland *J. Educ. Sociol.*, 34, No. 1, 7-16.

Hallworth, H. J. (1961). Teachers' personality ratings of high school pupils. *J. Educ. Psychol.*, 52, 297-302.

Morrison, A., McIntyre, D., and Sutherland, J. (1965). Teachers' personality assessments of primary school pupils. *Brit. J. Educ. Psychol.*, 35, No. 3, pp. 306-319.

Oliver, R. A. C., and Butcher, H. J. (1962). Teachers' attitudes to education—the structure of educational attitudes. *Brit. J. Soc. Clin. Psychol.*, 1, 56-69.

(*Manuscript received* 22nd *September*, 1965)

WHAT CAN PERSONALITY AND MOTIVATION SOURCE TRAIT MEASUREMENTS ADD TO THE PREDICTION OF SCHOOL ACHIEVEMENT?

By R. B. CATTELL* (*University of Illinois*)

A. P. SEALY (*London University*)

AND A. B. SWENEY (*Texas Technical University*)

SUMMARY. Intelligence (fluid-culture fair, and crystallized-traditional tests), fourteen primary personality factors (on the HSPQ), and fifteen dynamic structure factors (on the School Motivation Analysis Test), were measured on five samples of seventh and eighth grade children (Total N=563) involving (with the achievement test), eight hours of testing.

Each of the three modalities of predictor accounted for about a third of the criterion variance, and since each has its own space, they together accounted for almost three-quarters of the observable variance, leaving one-quarter for environmental influences and accidents outside the child. The nature of the personality and motivation correlations, e.g., positive relations with super ego and self sentiment factors, negatively with sensuality and pugnacity ergs, fits general motivation theory. Counsellers could thus increase predictive power by adding personality and motivation measures to the current ability tests.

I.—CONCEPTS AND INSTRUMENTS FOR PERSONALITY-MOTIVATION MEASUREMENT.

THE first half of this century saw a remarkable growth in the systematic understanding of human abilities, and of their predictive relation to various kinds of achievement. It produced a technology of scholastic measurement in which many thousands of specialists are now engaged and which has greatly increased the success of pupil placement, of remedial work with backward individuals, and of adaptation of curricula. It may well be that the second half will be a parallel development in the use of measures of personality and motivation in predicting and understanding achievement.

The basic problem is the lack of valid and meaningful measures of personality and motivational variables. It is possible to start with reasonably reliable scales measuring variables thought to be relevant in a particular situation. However, it will be difficult to interpret predictions of criteria, such as school achievement, obtained in this way unless the validity of the test is fully established, i.e., unless we know what the scale measures, and how this variable relates to other known variables. Such information is best obtained from comprehensive and systematic factorial studies of the structure of human personality and motivation, much in the way that Burt, Vernon, Thurstone and Guildford have studied human abilities.

In the last decade, basic factor analytic research on general personality and motivation structure has reached a degree of stability of replication which has justified setting up personality scales in terms of the objective tests (the O–A Batteries: Cattell, 1955) and questionnaires. The published versions of the latter at present available are: the High School Personality Questionnaire

---

* The research reported therein was carried out through the co-operative research program of the Office of Education, U.S. Department of Health, Education and Welfare.

(H.S.P.Q.) (Cattell and Beloff, 1953 ; Cattell, Karson and Nuttall, 1966), the Child Personality Questionnaire (C.P.Q.) (Porter, Schaie and Cattell, 1965) and the Early School Personality Questionnaire (Coan, Baker and Cattell, 1965). They measure general personality factors found to persist across age levels and across cultures (Tsujioka and Cattell, 1965 ; Cattell, Pichot and Rennes, 1961). These factors seem to be psychologically and clinically meaningful, including such structures as ego strength, schizothymia of temperament, super ego strength, surgency of disposition, etc., which have now been related to many criteria of real life performance. There are also relations mainly of primary to second order factors, between these Q-data (i.e., Questionnaire) factors and the T-data (behavioral test) personality factors on the O-A Battery (Hundleby, Pawlik, and Cattell, 1964). The latter have already been found significantly related to achievement (Pawlik, 1961) but the present article covers, in personality, only Q-data results.

The interest-motivation measures here to be related to achievement are those of the School Motivation Analysis Test (henceforth S.M.A.T.) (Sweney and Cattell, 1966). This test differs from such measures as the Strong Interest Test in that it measures unitary or independent motivation factors, derived from extensive and repeated factor analysis of interests, and provisionally identified as drives (ergs), e.g., fear, assertion, sex, etc., and sentiment structures. It differs from the Kuder factored test in measuring by *objective* or indirect methods such as information, word association memory, autism, etc. (Cattell, Radcliffe and Sweney, 1963) whereas the Kuder test and the Edwards' Personal Preference Scale use self-evaluative, verbal preference statements. The reader must be referred elsewhere for the extensive literature on the development of S.M.A.T. from concepts and experiments derived from the study of motivation measurement (Cattell, 1957, 1959 ; Cattell and Horn, 1963). The essential outcome of this break from the tradition of interest tests formerly in use is that (*a*) tension levels are measured in terms of unitary primary drives (strictly, ergs), and acquired sentiments such as sentiment to family, school, and self, and (*b*) the measurement is, for each drive, derived from a battery of diverse objective subtests sharing only demonstrated validity against the two major motivation component factors (Cattell, Radcliffe and Sewney, 1963).

II.—THE DESIGN, THE SUBJECTS AND THE PARTICULARS OF EXPERIMENTAL PROCEDURES.

The researches reported here seek to discover the direct relations between personality and motivation traits and level of achievement measured in a given year, that is to say, they deal with static individual difference relationships, although development analyses, e.g., concerning the relationships of change of achievement and change of motivation, have also been made from the same kinds of data elsewhere (Cattell and Sealy, 1965).

Between 1958 and 1964, there have been some five independent and extensive experiments investigating the relationship of structured personality and motivation traits to achievement. However, apart from Butcher's paper (Butcher, Ainsworth and Nesbitt, 1963) many of these studies have appeared only in Office of Education Reports (Cattell, Butcher, Connor, Sweney and Tsujioka, 1962 ; Cattell and Sealy, 1965) with limited circulation, and in previously unpublished papers given at A.P.A. meetings (Connor, 1962 ; Sealy, 1963), it seems desirable to survey the results together in a single article.

As pointed out, the correlational design is the same in all five studies. For the sake of cumulative knowledge, we also kept the measured variables essentially the same, as shown in Table 1.

TABLE 1

THE THIRTY-TWO SOURCE TRAITS MEASURED AND THE BATTERIES FOR THEM.

1. A Standard Scholastic Achievement Test.
   (Time, 3 hours).

2. An Intelligence Test:
   (*a*) Traditional, of "crystallized general ability."
   (Time, 60 minutes).
   (*b*) Culture Fair, "fluid general ability".
   (Time, 21 minutes) (Cattell, 1955).

3. The Fourteen factor High School Personality Questionnaire, Forms A and B.
   (Time, 80 minutes).
4. The Fifteen factor* School Motivation Analysis Test, Form A only.
   (Time, 70 minutes).

* This was an experimental version; the presently published S.M.A.T. is reduced to ten factors in line with M.A.T. (Cattell and Horn, 1964).

The H.S.P.Q. needs little description, and except for the listing of factor titles in Table 2, the reader is referred to its Handbook (Cattell, Karson and Nuttall, 1966) and the series of researches there listed. It will be noted that thirteen of the factors are the same as those in the Adult 16 P.F. above it, and the Child Personality Questionnaire (Porter, Schaie and Cattell, 1965) for the age range immediately below it. Intelligence has actually scored in two ways: (1) the fluid general ability factor measure (Cattell Culture Fair Tests), and (2) in an untimed, 'power' scoring of the crystallized ability test.

The School Motivation Analysis Test (Sweney and Cattell, 1966) not yet having been made available, deserves a little more description. The fifteen basic dynamic structure factors therein are derived from an interlocking series of factor analyses designed to cover the total interest sphere of children. With adults and children a number of factors have been replicated, and identified as ergs, such as sex and curiosity, or sentiments, such as the self sentiment or the religious sentiment (see Table 3). Each factor is scored as the sum of two 'attitudes' or interests found to load the factor most highly, which can be seen in the Handbook (Sweney and Cattell, 1966). As briefly indicated above, the fifteen factors are regarded as reflecting basic biological ergs ('drives'), on the one hand, and socially acquired patterns of behavior on the other.

The objective motivation measurement—or 'indirect' measuring devices mentioned above, are the most valid and suitable of sixty principles and devices explored in determining motivation component structure (Cattell, Radcliffe and Sweney, 1963). They are memory, information, word association and autism. The first is based on the principle that the greater an individual's interest in a course of action, the more efficient will be his perception and memory of verbal and pictorial material related to that course of action; the second illustrates that the greater his interest in something, other things being equal, the more he will know about it; the third, that he is more likely to choose word associations referring to something he is interested in, and the fourth test is based on the principle that a person is more likely to make 'autistic' judgments, or estimations, in the direction of his interests. The validities of these devices have been established by correlations with the motivation component

TABLE 2

POPULAR AND TECHNICAL TITLES OF THE FOURTEEN PERSONALITY FACTORS MEASURED BY THE H.S.P.Q.

Factor	Popular Title		Technical Title	
	Low Score	High Score	Low Score	High Score
A	Aloof, Stiff.	Warm, Sociable.	Sizothymia.	Affectothymia.
B	Dull Intellectually.	Bright Intellectually.	Low General Mental Capacity.	General Intelligence.
C	Emotional, Immature, Unstable	Mature, Calm.	Ego Weakness or Emotional Instability.	Ego Strength.
D	Stodgy.	Unrestrained.	Phlegmatic Temperament.	Excitability.
E	" Milk-toast," mild.	Aggressive, Competitive.	Submissiveness.	Dominance or Ascendance.
F	Sober, Serious.	Enthusiastic, Happy-go-Lucky.	Desurgency.	Surgency.
G	Casual, Undependable.	Conscientious, Persistent.	Lack of Acceptance of Group Moral Standards.	High Super Ego Strength.
H	Shy, Timid, Threat-Sensitive.	Adventurous, " Thick-skinned.	Threctia.	Parmia.
I	Tough, Realistic.	Esthetically Sensitive.	Harria.	Premsia.
J	Liking Group Action.	Fastidious Individualism.	Zeppia.	Coasthenia.
O	Confident, Self-Secure.	Timid, Insecure.	Confident Adequacy.	Guilt Proneness.
Q_2	Sociably Group Dependent.	Self-Sufficient, Resourceful.	Group Dependency	Self Sufficiency.
Q_3	Uncontrolled, Lax.	Controlled, Exacting Will Power.	Poor Self Sentiment Formation.	High Strength of Self Sentiment.
Q_4	Relaxed, Composed	Tense, Excitable, Frustrated.	Low Ergic Tension.	High Ergic Tension (General frustration).

factors measured (Cattell, Radcliffe and Sweney, 1963). These four test devices pair off into two motivational second order components: the ***integrated***, conscious and organized component and the unintegrated, less conscious and more hedonic component. (For meaning, see Cattell, Radcliffe and Sweney, 1963). These motivational components are, in fact, second order factors based on several factor analyses of the seven or eight primaries found in the above-mentioned wide array of sixty devices for measuring attitude or interests. Thus, for each area of interests (e.g., self-assertive interests), a measure becomes available of (*a*) the organized, integrated aspect of the interest, and (*b*) the less integrated, more hedonic and phantasy-level aspect of that interest. A list of the dynamic structure factors measured by these components is given in Table 3. The two second order motivation component factor scores, U (unintegrated) and I (integrated) can be added together if desired to give a total interest strength in the given erg or sentiment.

A special problem in objective motivation test use is that of eliminating the effect of the ' vehicles,' i.e., memory, reaction time, and other capacities of the individual through which interest is measured. This decontamination of

TABLE 3

NATURE OF THE 15 DYNAMIC TRAITS (ERGS AND SENTIMENTS) MEASURED BY S.M.A.T.

1.	Assertion	Ergs, measured on ergic tension levels.
2.	Sensuality	
3.	Sex	
4.	Gregariousness	
5.	Narcissism	
6.	Constructiveness	
7.	Protectiveness	
8.	Curiosity	
9.	Aggression	
10.	Play Phantasy	
11.	Acquisitiveness	
12.	Self Sentiment	Sentiments.
13.	Super Ego Sentiment	
14.	Religious Sentiment	
15.	Patriotic Sentiment	

interests from abilities, etc., has hitherto been accomplished by 'ipsatized scoring.' Therein the individual's total score on, say, the memory device, over *all* interests, is divided into each specific interest score, thus reducing all persons to the same total score through the memory device and eliminating differences due to memory power. Although this correction was used through the present researches, more recent developments in M.A.T. (Cattell and Horn, 1964) offer an improved alternative. For ipsatization gets rid both of differences in memory power (or other vehicles) and differences in total interest; and loss of the latter abolishes part of the true interest variance, thus making present interest-achievement correlations undoubtedly smaller than they naturally are.

For comparability, all five researches were done with children, not spread over college students or adults, and in the early adolescent 11–15 year (mainly 12–14 year) range. They were deliberately taken from both urban areas and rural areas, to give sociological and some geographical (change of school system) sampling. The five samples are shown in Table 4. Incidentally, Sample 5 is an independent experiment but not an independent study, as it consisted of samples 3 and 4 retested, on all variables, after a one-year delay (in pursuit of the developmental study) (Cattell and Sealy, 1965).

TABLE 4

THE FIVE EXPERIMENTS: SAMPLES OF CHILDREN STUDIED.

Study	Number of Subjects	Grade	Rural/Urban
1	154	7th and 8th	Urban
2	124	7th and 8th	Rural
3	61	8th	Urban
4	94	7th	Urban
5	130	8th and 9th	Urban*

* These are approximately the same samples of children as 3 and 4 tested one year later.

The total testing time for the battery described in Table 1 was eight hours, in almost all cases spread out over four days and the testing was carried out in a standard classroom situation.

III.—ANALYSIS OF RESULTS : FOR THE THREE MODALITIES IN ISOLATION.

As pointed out above, in this 'static' or 'absolute level' analysis (in contrast to the change or developmental study) we simply proposed to ask, as has been traditionally asked with intelligence and primary ability measures, what the correlation is between achievement and (*a*) the ability, (*b*) the personality-termperament, and (*c*) the dynamic modalities of source traits. Further interpretation, in causal terms, of dependent and independent variables, is postponed, though we are explicitly rejecting the *assumption* that personality, ability and motivation are simply and only antecedent variables, for there is clear evidence, e.g., in the work of Tsushima (Cattell and Scheier, 1961) that achievement experiences in turn affect personality and motivation.

First, the reliabilities of the variables to be related are shown in Table 5.

TABLE 5

RELIABILITY (DEPENDABILITY) COEFFICIENTS FOR VARIABLES INVOLVED.

Culture Fair Intelligence Test (Forms A and B) = 21 mins. = ·90.
Crystallized Intelligence Test (W.I.S.C) = 60 mins. = ·93.
Achievement Test (Metropolitan and California mean) = 90 mins. = ·90.

Personality Factors, HSPQ Forms A and B together : 5¼ mins. per factor		Dynamic Factors, SMAT Form A only : 2¼ mins. per factor (Ipsatized scores)		
		Trait	Integrated	Unintegrated
A	·76	Assertion	·51	·12
B	·74	Sensuality	·45	·25
C	·75	Sexuality	·49	·11
D	·79	Gregariousness	·30	·17
E	·71	Narcism	·51	·20
F	·69	Constructiveness	·32	·05
G	·70	Protectiveness	·49	·22
H	·81	Curiosity	·45	·25
I	·78	Pugnacity	·38	·18
J	·68	Play-Phantasy	·39	·14
O	·71	Acquisitiveness	·40	·08
Q_2	·71	Self Sentiment	·39	·14
Q_3	·72	Super Ego Sentiment	·34	·16
Q_4	·80	Religious Sentiment ..	·41	·16
		Patriotic Sentiment ..	·42	·15

These are the mean values across the two samples for which alone actual retest data exist, i.e., they are *dependability coefficients*.

The present reliabilities of intelligence, personality and achievement (the latter, here, obtained on only a sub-sample) are in accord with those usually found and if anything indicate slightly better than usual care on the part of the children. The dynamic, motivation measures are also reasonably good in relation to other experience, but show the usual far lower stability coefficients, especially for the unintegrated component. These lower reliabilities are customary with ipsative scoring, because much of the literal, observed score variance is removed, and one must be on guard against viewing these figures in terms of accustomed direct measurement reliabilities. However, as indicated in Table 5, some of the reduction of reliability is due to the brevity—2½ minutes—of the scales concerned. The arguments for proceeding with such brief scales and accompanying reduced reliabilities are (*a*) the practical one that with a total of seven to eight hours of testing time from the children, it is a question of short scales or none ; (*b*) the strategic consideration that these are exploratory studies, in which preliminary evidence only is sought on the role of motivation, and (*c*) that by correction for attenuation either in correlations or factor analyses averaged across four studies, we can, nevertheless, hope to obtain a good estimate of the true importance of the dynamic factors.

An analysis of correlations with the criterion is first made for ability and personality factors in Table 6. Although in practice one would urge school psychologists always to use both forms (A and B) of the H.S.P.Q. for adequate reliability, yet in examining the significance and consistency of results, it is helpful to see results for the single scales (provided one recognizes that the criterion relations will be automatically attenuated in magnitude thereby).

TABLE 6

CORRELATIONS BETWEEN PERSONALITY FACTORS (AND INTELLIGENCE) WITH SCHOLASTIC PERFORMANCE

Factors	Sample 1 N=154 Urban	Sample 2 N=124 Rural	Sample 3 N=61 Urban 7th Graders		Sample 4			Sample 5 N=130 Urban 8th and 9th Graders	
					N=94 Urban 8th Graders		N=155 Urban 7th and 8th Graders		
	7th and 8th Graders	7th and 8th Graders	Form A only	Form B only	Form A only	Form B only	Forms A and B Together	Form A only	Form B only
A+ Affectothymia	·11*	·10	—·04	·11	·10	·24*	·13	·09	·03
B+ Intelligence	·51†	·39†	·44†	·36†	·31†	·47*	·50†	·46†	·31†
C+ Ego Strength	·01	·00	—·17	—·20	·08	—·01	—·08	—·19*	—·12
D+ Excitability	—·07	·00	·31*	—·16	—·06	—·04	·02	·01	—·01
E+ Dominance	—·18*	—·18*	—·11	—·04	—·01	·15	·01	—·07	—·00
F+ Surgency	·02	—·06	·09	—·01	·10	·17*	·13*	·06	·01
G+ Super Ego Strength	·31*	·14	·13	·28*	·14	·20*	·21*	·19*	·09
H+ Parmia	—·01	·01	—·04	·06	·06	·08	·03	·06	—·12
I+ Premsia	·06	—·17	·09	·05	·08	—·01	·10	·10	·05
J+ Coasthenia	·03	·02	·30†	·05	—·08	·18*	·19*	·16*	—·01
O+ Guilt Proneness	—·03	—·08	·03	·03	·06	—·12	·01	·16*	—·01
Q$_2$+ Self Sufficiency	·25†	·10	—·05	—·11	—·14	—·16	—·19*	—·08	—·15*
Q$_3$+ Self Sentiment Strength	·13*	·07	·18	·25*	·19*	·24*	·24*	·09	·28†
Q$_4$+ Ergic Tension	—·02	—·02	·02	—·01	—·17*	·12	—·02	·20*	·07
Crystallized Intelligence (Calif. Mental M.T.)							·67†	Both Samples	·68†
Fluid Intelligence (Culture Fair Intelligence Test)	Both Samples	·46						Both Samples	·48

* Differs from zero with probability < ·05.

† Differs from zero with probability < ·01.

The achievement relation with the three measures of intelligence hovers in the +·5 to +·6 region to which we are accustomed in the middle school range. If we increase the length of the 16 P.F. intelligence (B factor) scale by a ratio of 10, to bring it to the same length as the California Mental Maturity Test, the two crystallized general ability measures approach each other, the 16 P.F. B standing at 0·55, very close to the central value of Thomson's Scottish survey and other painstaking and adequate studies and the Cal. M.M.T. at 0·68. The size of the latter suggests that this test is more contaminated with school skills, i.e., is predicting achievement from achievement, even more than most crystallized intelligence measures. The true role of intelligence—of fluid general ability—is indicated rather by the Culture Fair Test which, if also corrected to the standard one hour length (by a ratio of 3) becomes 0·50.

Turning next to the personality factors, we find several consistently significant relations, notably of affectothymia, A, positive (Samples 1 and 4) ; dominance, E, negative ; super ego strength, G, positive (Samples 1, 3, 4 and 5) ; coasthenia, J., positive (Samples 3 and 4), and self sentiment strength, Q_3, positive (Samples 1, 3, 4 and 5). The behavior of Q_2, self sufficiency, is peculiar, being significantly positive in one school system and a young age and negative in another and at older age. Q_2 like A and F are components of the higher order Extraversion factor, and there have been suggestions in previous work that these factors favour school progress in some samples of children and inhibit it in others ; these samples have varied in terms of age, quality of school, teaching methods (e.g., emphasizing group co-operation or individual effort), rural-urban location, etc., but no systematic variation has yet been found to account for these apparent reversals in the relationship. Ergic tension, Q_4, similarly presents ambiguous results. The absolute size of the A, E, G, J and Q_3 correlations, when a corrected best estimate is made, are shown in Table 7.

TABLE 7

SIGNIFICANT CORRELATIONS OF PERSONALITY FACTORS AND ACHIEVEMENT CORRECTED FOR ATTENUATION ON PERSONALITY MEASUREMENTS.
(Criterion Left As It Stands.)

Personality Source Trait		Mean of 8 Correlations* One Scale 2½ Mms. 10 Items	Correlation With Achievement Corrected for Attenuation
Affectothymia	A	·09	·12
Dominanace	E	—·06	—·09
Super Ego	G	·19	·26
Coasthenia	J	·08	·12
Self Sentiment	Q_3	·18	·24

* These values are worked out for factors with significant achievement correlation on at least two samples, but are averaged over *all* samples (for the single form values).

Turning next to the dynamic trait correlations from S.M.A.T., we see in Table 8 that, in spite of the only moderate reliabilities the number of significant criterion relations is appreciable. Some difficulty of comparison arises because the dynamic traits at the time of the first two samples were measured on the total score only, and we since have learnt that this may lose significant relations, because the I and U components are, in general, uncorrelated (Cattell, 1957) and may, sometimes, correlate significantly in opposite directions with a criterion. Thus, with the present criterion, achievement, the self assertive drive is significantly correlated positively in its integrated components and negatively in its unintegrated expression. There are strong indications of the same with sensuality, sex and gregariousness, and in all four of these ergs, the summed single total drive score, neglecting components, has near zero relation to the criterion.

TABLE 8

DYNAMIC TRAITS RELATED TO SCHOOL ACHIEVEMENT.

Dynamic Trait (Erg or Sentiment)	Measured in Total Motivation Strength : U+I		Measured on Integrated, I, and Unintegrated U, Components Separately					
	Sample 1 N= 154 Urban 7th and 8th Graders	Sample 2 N= 124 Rural 7th and 8th Graders	Sample 3 N=61 Urban 7th Graders		Sample 4 N=94 Urban 8th Graders		Sample 5 N=130 Urban 8th and 9th Graders	
			U	I	U	I	U	I
1. Assertion	–·02	–·06	·01	·33†	–·20*	·11	–·15*	·09
2. Sensuality	–·05	·06	–·13	–·17	·25*	–·16	·08	–·46†
3. Sex	–·02	–·07	–·23*	·14	·02	·03	–·06	·26*
4. Gregariousness	–·02	–·01	–·04	·09	·08	–·22*	·18*	·27*
5. Narcissism	·06	·09	·17	–·16	–·08	·06	–·06	–·06
6. Constructiveness	·04	·08	–·05	–·12	·01	–·13	–·18*	–·21*
7. Protectiveness	–·06	–·16*	–·02	–·21*	–·11	–·18*	–·34†	·02
8. Curiosity	–·12*	–·06	·14	–·21*	·07	–·07	·06	·16
9. Pugnacity	–·28†	–·09	–·12	·01	–·04	·01	–·16	–·14
10. Play Phantasy	·02	–·13	·19	·03	–·03	·01	·07	–·01
11. Acquisitiveness	·11*	–·06	·04	·09	–·08	·09	·11	·09
12. Self Sentiment	·06	·04	·20	·21*	·14	·16*	·03	·06
13. Super Ego Sentiment	·06	·12	·06	–·08	–·17*	·05	–·02	·04
14. Religious Sentiment	·07	·11	–·06	–·08	–·04	·04	·04	·03
15. Patriotic Sentiment	—	—	·07	·15	–·04	·27*	·13	·11

Beyond these, the significant correlations with achievement are negatively with Pugnacity (more from the unintegrated component) and with Curiosity, positively with the Self Sentiment, with Acquisitiveness, and with Patriotic Sentiment, and negatively with Parental Protectiveness and Constructiveness. These will be discussed in the summary. Incidentally, it will be noted that drive strengths are slightly more negatively than positively related to achievement. (There are fifty-six negative and sixty-two positive correlations; sixteen significantly negative, six significantly positive). This may be surprising if we think of total drive, like total general ability, as being conducive to higher performance generally. But it must be remembered that ipsative scoring has got rid of the general drive factor, and some of the negative trend in correlation may be due to a slight negative correlation produced with the vehicles, which are abilities (Cattell and Horn, 1964). It is almost certain that if dynamic traits were scored without ipsatization (as in some use of the M.A.T., Cattell and Horn, 1964) more positive correlations would become significant.

To estimate what the magnitudes of dynamic trait predictions of achievement will probably reach when longer, more reliable scales are introduced, let us consider only those ergs and sentiments for which there is cumulative evidence of significance of correlation and correct them for attenuation as in Table 9.

TABLE 9

SIGNIFICANT CORRELATIONS ON DYNAMIC FACTOR AND ACHIEVEMENT, CORRECTED FOR ATTENUATION ON DYNAMIC MEASURES.
(Criterion Left As It Stands.)

Dynamic Source Trait*	Mean r Across Samples	Mean r Corrected for Attenuation
Assertion, Unintegrated	—11	—·30
Assertion, Integrated	18	·26
Sensuality, Integrated	—26	—·39
Gregariousness, Integrated	—13	—·24
Protectiveness, Unintegrated	—16	—·35
Protectiveness, Integrated	—12	—·17
Pugnacity, Unintegrated and Integrated	—28	—·45
Self Sentiment, Integrated	15	—·25

* The dynamic traits here are those for which there is at least one r at the P< ·01 level or r's on two samples at the P< ·05 level Their r's are then averaged across all samples The test reliabilities are from Table 5 above.

These correlations are of the same order as those obtained by Cattell, Butcher, Connor, Sweney and Tsujioka (1962) by the alternate route of factoring and rotating the pure dynamic factors to simple structure in relation to the criterion factor.

IV.—ANALYSIS OF RESULTS: FOR MULTIPLE PREDICTION RELATIONSHIPS.

In the report of the first experiment in this area, we emerged with the conclusion that the ability modality, the personality-temperament modality and the dynamic, interest modality probably contributed about equally to achievement prediction, and that since they were substantially independent sets of factors, one might expect to increase the customary prediction of achievement variance from ability test by close to 200 per cent. (Cattell, Butcher, Connor, Sweney, and Tsujioka, 1962).

Both for practical educational purposes of maximizing prediction, and for the theoretical psychological purposes of understanding trait interaction, it is desirable now to look at the degree of independence of the source traits in the different modalities, and to calculate the multiple correlation (and its shrinkage) that may be possible from using ability, personality and dynamic trait (interest) measures.

A theoretical, operationally-expressed basis of clear distinction among the concepts of ability, temperament and motivation, has been made (Cattell, 1946), which would argue for their being essentially uncorrelated. Empirical evidence to date has shown very little correlation of ability and general personality factors (Cattell, 1957 ; Cattell and Scheier, 1961) in the second order, but some correlation between general personality factors and dynamic, motivation factors (Cattell, 1957, page 527). In two instances—the self sentiment and the super ego strength factors—as measured respectively by personality questionnaires and objective motivation tests, we are almost certainly measuring the same factors. But there is extensive evidence (Cattell, Butcher, Connor, Sweney, and Tsujioka, 1962) that the ergs and the ' non-personal ' sentiment factors are occupying and describing a realm largely beyond the realm of known ability and temperament factors.

Faced with sets of dimensions in three spaces, but with some minor overlap in common space, our representation of results can proceed in three main ways : (1) By calculating the multiple correlation first for abilities, secondly for ability plus personality, and thirdly for ability, personality and motivation all together. The increments will not represent the full contribution of the later modalities, because whatever is in common space will have been credited to the earlier ones. There are six possible orders in which this could be done, but we shall handle this in the traditional manner by taking the established predictors first, that is, ability measures, and the newest of class predictors (i.e., motivation measures last). (2) Secondly, and more fairly, we can calculate the multiple R separately for each modality and compare these values. (3) We can remove the common factor variance and examine the prediction peculiar to what remains with each modality. The last is a somewhat artificial procedure. The fact is that some real though slight overlap exists when the three modalities are each correctly measured, just as with oblique factors. At the present stage of research, it seems best simply to examine the relations of the modalities by methods (1) and (2).

In regard to the amount of achievement predicted by intelligence, the recent findings of Horn (1964) and of Cattell (1963) have suggested the need for a distinction between the fluid general ability factor and the crystallized general ability factor. The traditional intelligence test, e.g., the Binet, the Wechsler, etc., is not measuring a single general ability in the Spearman sense (Cohen, 1957) but two general ability factors, g_f (fluid ability) and g (crystallized ability) and the latter is substantially an expression of school achievement skills and similar to Vernon's *v. ed.* factor (Vernon, 1961). As the above correlation of 0·68 of achievement with the California test shows, the latter is particularly heavily contaminated with school attainment. If we think—as some designers evidently do—that the best intelligence test is one which maximally predicts school achievement, then let us frankly make the test as close as possible to a measure of scholastic skills and information, for one achievement test is the best predictor of another. But, if, as psychologists, we wish to analyze the causes of achievement, in general and in terms of understanding the individual case of backwardness, it is desirable to separate out the functionally unitary core of intelligence which is the fluid general ability factor, g_f, as measured by Culture Fair Intelligence Tests. The tendency in the achievement correlation of good traditional intelligence tests has been to center upon a value of +0·55 (with which the present value of the California test is out of line). As Table 10 (a) reminds us, the Culture Fair value here is ·46. Probably with various culture fair tests corrected for attenuation to a standard one-hour length the r at this school age would be about +·50 (criterion uncorrected, but of three-hour length). It seems that about 25 per cent. of the variance of school achievement is predictable from general intelligence.

In Table 10 we present the multiple R's first for samples 1 and 2 combined (278 cases) and separately for samples 3 and 4 combined (144 cases). (For data in the latter are in a form to permit finer breakdowns, as we there have *two* achievement measures—school classroom grades *and* a standard achievement test (Metropolitan and Californian).

TABLE 10

MULTIPLE PREDICTION OF ACHIEVEMENT FROM INTELLIGENCE, PERSONALITY AND DYNAMIC TRAITS IN VARIOUS COMBINATIONS.

(*a*) Samples 1 and 2, Urban and Rural, 7th and 8th Graders. N=278.

Measure	Number of Variables	Multiple R With Achievement	Corrected for Shrinkage	Percentage of Achievement Variance Predicted
1. Culture Fair (Fluid Intelligence)	1	·46	—	21%
2. Personality Factors, HSPQ Forms A and B (Without Intelligence Scale)	13	·55	·48	30%
3. Intelligence, plus Personality, plus Dynamic Traits..................	29	·72	·60	52%

(*b*) Samples 3 and 4, Urban, 7th and 8th Graders. N=144*.

Measure	No. of Variables	Multiple R With Achievement		Corrected for Shrinkage		Percentage of Achievement Variance Predicted	
		(a) Standard Achiev	(b) Grades	(a) S.A.	(b) Gr.	(a) Standard Achiev	(b) Grades
1. Culture Fair (Fluid Intelligence)	1	·50	·48	—	—	25%	23%
2. Personality Factors, HSPQ Forms A and B. (Without Intelligence Scale)	13	·52	·60	·33	·46	27%	36%
3. Dynamic Traits (SMAT Unintregated)	15	·34	·37	·11	·19	11%	14%
Dynamic Traits (SMAT Integrated)	15	·40	·32	·24	·50	16%	10%
Dynamic Traits (SMAT U and I Together)	30	·52	·48	·28	·16	27%	23%
(1+2) Intelligence and Personality Factors	14	·79	·75	·73	·68	62%	56%
(1+3) Intelligence and Dynamic Traits (U and I)..............	31	·77	·67	·70	·54	59%	45%
(2+3) Personality and Dynamic Traits (U and I)	18	·69	·72	·37	·45	49%	52%
(1+2+3) Intelligence, Personality and Dynamic Traits (U a d I)..	32	·85	·81	·73	·65	72%	66%

* Eleven children were dropped from the group in Table 6 because of incomplete data for the multiple relations.

The variance contributions for the separate modalities are thus 21 to 25 per cent. for intelligence, 27 to 36 per cent. for personality traits, and 23 to 27 per cent. for dynamic interest traits. It may be said that they are roughly on a par. The combinations two at a time are also roughly on a par, in the 45 to 62 per cent. range. In accordance with our previous findings and theory (Cattell, 1957) of slightly greater overlap of personality and motivation than of other modality pairs this combination gains a little less than the others ; but the general position is sustained, by the essential additiveness of the variance contributions, that is, the three modalities are making largely separate contributions.

Finally, when all three are brought together we reach a multiple R of +0·85, possibly shrinking on a new sample to +0·73. Incidentally, except for personality predictors, there is a slight tendency for predictions to run lower for school grades than for results of the external standard examination. There is no way of obtaining reliabilities for grades, but they probably have somewhat lower reliability than the standard examination. If this is so, the fraction of the criterion predicted by these three modalities is in the range of 70 to 75 per cent.—the upper limit with corrections for attenuation.

The difference between the +0·85 multiple correlation reached in the second pooled sample (N=144) and that of +0·72 in the first (N=278) could be partly a sampling variation, but is almost certainly due to the relatively crude form of S.M.A.T., in the pioneer research. For the intelligence and personality tests used were identical, and their variance contributions to the criterion turn out to be essentially the same.

V.—SUMMARY AND DISCUSSION.

(1) A culture fair intelligence test, the H.S.P.Q. measuring thirteen, personality factors and crystallized intelligence (traditional intelligence test) and the School Motivation Analysis Test (S.M.A.T.) measuring fifteen dynamic interest traits have been administered to five samples (N=154, 124, 61, 94, and 130) of seventh and eighth grade (13 to 15-year-old) boys and girls, for whom school class grades and a standard achievement test scores were also available. The results represent approximately eight hours of testing under standard classroom conditions.

(2) The correlations of the two kinds of achievement measurement with these predictors are closely parallel. They confirm the main position taken up from the first pioneer study (Cattell, Butcher, Connor, Sweney and Tsujioka, 1962) that (*a*) the three modalities of traits—ability, temperament and interest—are essentially in separate spaces, with only minor intercorrelations, and (*b*) the approximate formula that one-quarter of the variance of achievement is associated with abilities, one quarter with personality differences, one-quarter with dynamic interest differences, and the remaining quarter with influences and accidents outside the individual.

(3) Beyond this main conclusion, there is some indication of the need for further research, which may qualify it, and which requires closer attention to variations in test reliabilities, school systems and cultures, ages of subjects, etc. Thus, at present there are indications that personality contributes slightly more than its third and motivation slightly less. But the measurement of dynamic traits by objective (non-questionnaire) methods in S.M.A.T., is a new venture, and the reliabilities with ipsative scoring are low. Results of correction for

attenuation indicate that achievement prediction from this modality, with lengthened and improved tests *could* exceed that from either of the other modalities. Research is, therefore, now needed in (*a*) other schools, ages and areas, in which (*b*) reliabilities for the criteria are also introduced into the calculations, (*c*) lengthened H.S.P.Q. and S.M.A.T. tests (by additional forms) are employed, (*d*) the nature of the small overlap of the modalities is clarified in terms of personality structure, (*e*) some additional source trait dimensions are introduced where known to exist, and (*f*) the best possible general specification equation and its variation and shrinkage from sample to sample are examined.

(4) In terms of causal explanation we have not committed ourselves to the position that the associations found represent causal action by the predictors. Sequential studies, of which two or three exist (Tsushima, *op. cit.*, Cattell and Sealy, in press) show that achievement may affect personality and motivation. They also show (Cattell and Sealy, in press) that *current* motivation change is more highly related to achievement than absolute motivation level. Achievement is obviously the cumulative result of ability, personality and motivation operating over many years. Personality and motivation, especially are both subject to change, to a greater degree than we are accustomed to in ability measures, and a model from integral calculus would be necessary adequately to relate personality and motivation scores, over years, to achievement. The present simple, static additive model falls far short of what is theoretically required, thus the fact that these measures at a given year, predict between a half and three-quarters of the variance (allowing for shrinkage) is remarkable.

(5) From the standpoint of personality theory the actual traits found significantly associated with school achievement are readily explicable and their relation adds new facets to their meaning as traits. In dynamic traits, the self assertive erg is correlated positively in its integrated form, as one would expect, but negatively in the unintegrated component on current theory, a higher score on unintegrated then integrated component would only occur when the individual is failing to satisfy his self assertion and is seeking substitute satisfaction in fantasy. The negative association between pugnacity and achievement may imply the existence of frustration reactions to failure experiences ; the negative relationships with sex, sensuality and gregariousness may suggest their role as distractors from school work. What remains obscure, despite its consistency over the samples, is that parental, protective tender concern should be negatively related, and that this small negative relation persists when the sex of the subject is partialled out. Is a lack of altruistic concern for others, perhaps facilitating to scholastic concentration and success ? That the self sentiment, as an integrator of motives, and the acquisitive erg, devoted here to the collection of grades, should be positively related, fits our other evidence on their nature.

Among personality source traits, the associations are in agreement with those found in other studies (Pierson, *et al*, 1964), namely, positively with affectothymia, A, negatively with dominance (independence), E, and positively with super ego strength, G, and self sentiment strength, Q_3. There is some failure of agreement in (*a*) finding J related to achievement and (*b*) not finding Q_2 related. Reasons for the G. and Q_3 relations are obvious. E is negatively related because the independence and lack of docility obstruct classroom learning. Why affectothymia should be positively related in school though negatively at the later college level (a more understandable liking of the A— person for books) is something of a mystery.

The ultimate aim of this line of research must be to write a specification equation on the following lines:

$$\text{Achievement} = \cdot 1A + \cdot 5B - \cdot 1E + \cdot 3G + \cdot 1J + \cdot 2Q_2 + \cdot 3Q_3 - \cdot 3As(U) + \cdot 2As(I) - $$
$$= \cdot 4Sn(I) - \cdot 2Gr(I) - \cdot 3Pr(U) - \cdot 4Pu$$

(where A, B, etc., are the personality and ability factors and As, Sn, etc., are the dynamic structure traits of assertiveness, sensuality, etc.).

Such an analysis aided by psychological as well as statistical laws, will be susceptible of modification to other groups and to understanding the individual case.

It will be obvious from this specification equation and the number of factors involved that (*a*) we should not expect, with essentially independent trait factors that correlations, though significant, will run in general above ·1 and ·2. In other words, these low values are not due to test unreliability, but express the reality of the complex determination of achievement by personality traits;

(*b*) these are consistent with our claim that the prediction of level educational achievement from ability tests could be doubled in accuracy by adding personality source traits and trebled by adding objectively measured dynamic traits. It seems likely that 75 per cent. predictions of the criterion variance is the limit for these types of measures, since they take no account of variations in school, teaching quality, home facilities, or the intervention of illness and accidents in the examination room. The encouraging magnitude of this prediction does, however, justify our position that educational prediction should not be based on *ad hoc* tests and scales, but on comprehensive tests of the main source traits in personality structure and the main dynamic traits in the total interest area.

VI.—REFERENCES.

BUTCHER, J., AINSWORTH, M., and NESBITT, J. E. (1963). Personality factors and school achievement. *Brit. J. Educ. Psychol.*, Vol. 33, 276-285.

CATTELL, R. B. (1946). *The Description and Measurement of Personality.* New York: World Book Co.

CATTELL, R. B. (1957). *Personality and Motivation Structure and Measurement.* New York: Harcourt, Brace and World.

CATTELL, R. B. (1959). The dynamic calculus: concepts and crucial experiments. *Nebraska Symposium on Motivation.* Univer. of Nebraska Press.

CATTELL, R. B. (1963). Theory of fluid and crystallized intelligence: a critical experiment. *J. Educ. Psychol.*, 54, 1-22.

CATTELL, R. B., BAGGALEY, A. R., CHECOV, L., COGAN, E. A., FLINT, D., GRUEN, W., HUSEK, E., MEELAND, T., SAUNDERS, D. T., and SCHIFF, H. (1955). *Handbook for the Objective-Analytic Personality Test Batteries.* Champaign, Ill.: The Instit. for Pers. and Abil. Testing.

CATTELL, R. B., and BELOFF, H. (1953). Research origin and construction of the I.P.A.T. Junior Personality Quiz. *J. Consult. Psychol*, 17, 436-442.

CATTELL, R. B., BUTCHER, J., CONNOR, D., SWENEY, A. B., and TSUJIOKA, B. (1962). *Prediction and Understanding of the Effect of Children's Interest upon School Performance.* Urbana, Illinois: Lab. of Pers. Assess. and Grp. Behav., University of Illinois.

CATTELL, R. B., and CATTELL, A. K. S. (1955). *The Culture Fair Intelligence Test, Scale* 2. Champaign, Illinois: The Instit. Pers. and Abil. Testing.

CATTELL, R. B., and HORN, J. (1963). An integrating study of the factor structure of adult attitude-interests. *Genet. Psychol., Monogr.*, 67, 89-149.

CATTELL, R. B., and HORN, J. (1964). *The Motivation Analysis Test.* Champaign, Illinios. Instit. Pers. and Abil. Testing.

Cattell, R. B., Karson, S., and Nuttall, R. *Handbook for the High School Personality Questionnaire*. Champaign, Ill. : Instit. Pers. and Abil. Testing (in press).

Cattell, R. B., Pichot, P., and Rennes, P. (1961). Constance interculturelle des facteurs de personnalite mesures par le test 16 P.F. II. Comparaison franco-americaine. *Revue de Psychologie Appliquee*, 11, 165-196.

Cattell, R. B., Radcliffe, J. A., and Sweeney, A. B. (1963). The nature and measurement of components of motivation. *Genet. Psychol., Monogr.*, 69, 49-211.

Cattell, R. B., and Scheier, I. H. (1961). *The Meaning and Measurement of Neuroticism and anxiety*. New York : Ronald Press.

Cattell, R. B., and Sealy, A. P. (1965). *The General Relations of Changes in Personality and Interest to Changes in School Performance : An exploratory Study*. Report to the Office of Education on Co-operative Research Program Project No. 1411. Laboratory of Personality Assess., Univer. of Illinois.

Cattell, R. B., and Sealy, A. P. (1966). *The Trends in Adult Personality Source Traits in our Culture, as Measured by the* 16 *P.F. Test* (in press).

Cattell, R. B., and Sweney, A. B. (1964). Components measurable in manifestations of mental conflict. *J. Abn. Soc. Psychol.*, (in press).

Cattell, R. B., Sweney, A. B., and Sealy, A. P. (1964). *School Motivation Analysis Test Handbook* (in press).

Coan, R. W., Baker, R., and Cattell, R. B. (1965). *The Early School Personality Questionnaire*. Champaign, Illinois : The Instit. for Pers. and Abil. Testing.

Cohen, J. (1957). The factorial structure of the W.A.I.S. between early adulthood and old age. *J. Consult. Psychol.*, 21, 283-290.

Connor, D. V. (1962). *Predicting School Achievement by Interest and Motivation Tests.* Paper given to the Annual Conference of the A.P.A., August, 1962.

Horn, J. (1964). *A Test of the Theory of Fluid and Crystallized General Intelligence.* Doctoral dissertation, Univer. of Ill. Library, Urbana, Illinois.

Hundleby, J. D., Pawlik, K., and Cattell, R. B. (1964). *Personality Factors in Objective Test Devices*. San Deigo : Knapp.

Pawlik, K. (1900) Educational prediction from objective personality test dimensions. *Amer. Psychol., Assoc , Inc.*, 06, 6, 422.

Pierson, G. R., Barton, V., and Hey, G. (1964). S.M.A.T. motivation factors as predictors of academic achievement of delinquent boys. *J. Psychol.*, 57, 243-249.

Porter, R. B., Schaie, W., and Cattell, R. B. (1965). *Child Personality Questionnaire Handbook*. Champaign, Illinois : Instit. of Pers. and Abil. Testing, 2nd Edit.

Sealy, A. P. (1963). The influence of children's interests and personality upon their scholastic achievement. *Amer. Psychol. Assoc., Inc.*, 18, No. 7.

Sweney, A. B., and Cattell, R. B. (1961). Dynamic factors in twelve-year-old children as revealed in measures of integrated motivation. *J. Clin. Psychol.*, 17, 360-369.

Sweney, A. B., and Cattell, R. B. (1966). *The School Motivation Analysis Test.* Champaing, Illinois : Instit. Pers. and Abil. Testing.

Tsujioka, B., and Cattell, R. B. (1966). *Constancy and Difference in Personality Structure and Mean Profile, in the Questionnaire Medium, from Applying the* 16 *P.F. Test in America and Japan* (in press).

Tsushima, Y. (1957). *Failure Stress in Examinations Related to Anxiety Inventory Scores.* Masters thesis, Univer. of Illinois Library, Urgana, Illinois.

Vernon, P. E. (1961). *The Structure of Human Abilities.* London : Methuen.

(*Manuscript received 9th September*, 1965)

THE ROLE OF ERRORS IN LEARNING WITH FEEDBACK

By WARWICK B. ELLEY

(*University of Auckland*)

SUMMARY. The extent to which errors interfere with efficient learning was examined in two experiments in which punchboards were used to provide immediate feedback. Rote and logical learning tasks were learned under varying and deliberately manipulated error rates. As predicted, rote learning efficiency was reduced under high error rates, and error repetition was considerable. With meaningful, non-arbitrary tasks, however, initial error rates were unrelated to end-test performance, and error repetition was uncommon. Implications were drawn for programmers and teachers.

I.—INTRODUCTION.

THE extent to which errors interfere with learning has been recurrently debated in the literature on programmed instruction. Skinner (1958) has supported a low error rate of 10 per cent. or less per frame, arguing that error tendencies once made tend to persist. Crowder (1963) contended, on the other hand, that it is impractical and undesirable to make programs so easy that errors are eliminated. Williams (1963) has summarised the debate by listing the logical arguments for and against minimal error. He concluded that there are probably optimal error rates for program frames, but that these would vary according to program content and learner characteristics.

Unfortunately, little is known about the relationship between error rates and the characteristics of learners or subject matter. Indeed, insufficient verification has been made of Skinner's contention that incorrect responses made to multiple-choice questions would interfere with learning at all. Thorndike (1932) did demonstrate that errors made in rote learning tasks, such as the learning of Spanish vocabulary, do tend to persist. In fact, as a result of such experiments, he was persuaded to revise his law of effect by denying the weakening effect of dissatisfaction on connections already formed. Nevertheless, the statistical problems he faced in establishing a base-line from which to judge the extent of connection weakening, have thrown suspicion on his conclusions (Hilgard, 1956), while the fact that feedback in his experiments was limited to a simple announcement of ' Right ' or ' Wrong ' makes extrapolation to programming a dangerous practice. The typical program both confirms and corrects, making error repetition seem less likely.

Kaess and Zeaman (1960) altered Thorndike's design by using a punchboard device to provide more informative feedback and by varying the number of plausible distractors faced by each group in the learning of psychological terms. Again *S*s tended to repeat their errors. Faced with five choices per item on the punchboard *S*s made more errors in an end-test than those with fewer distractors. Moreover, wrongly punched answers were shown to be more attractive for *S*s of the two-choice condition, than were new distractors in the end test. This study avoided Thorndike's statistical problem by randomly assigning the single wrong distractor in the two-choice items, thus allowing for an objective determination of the chance base line in assessing the attractiveness of each option in the end-test. Skinner's contention stands supported by this experiment. Apparently plausible wrong answers, once selected, tend to recur, even though they are immediately corrected.

There is a further difficulty in extrapolating from these experiments, however. All were conducted with rote learning tasks, where arbitrary connections were to be set up by the learner. There is now sufficient evidence (Ausubel, 1963) on the qualitative differences between rote and meaningful learning that one is justified in questioning the extent to which Skinner's minimal error-rate principle would apply in the latter. The popular belief that we profit from our mistakes makes better sense in learning tasks where a meaningful non-arbitrary pattern can be discerned by the learner. And it is with this kind of learning that most program writers are predominantly concerned. Unfortunately, the principle of minimal error is difficult to test in a programming context, since the deliberate reduction of error by program alteration must usually be associated with other contaminating factors, such as a variation in the length of the frame or the program. The principle can be tested with multiple-choice questions, however, by varying the number of plausible choices offered the learners, as was done by Kaess and Zeaman (1960). By this procedure, the amount of information presented and the number of reinforced responses is constant across groups, while the number of errors made depends more on the experimental manipulation than on the ability of the learners.

Two experiments were designed to test the hypothesis that errors play different roles in rote and meaningful learning under conditions of feedback presented by a punchboard. It was predicted that low error rates would facilitate learning in rote but not meaningful tasks.

II.—EXPERIMENT 1.

Method.

Subjects.—*S*s were sixty-six second-year teachers' college students from five heterogeneous classes. All were naive to punchboard learning tasks.

Materials.—For the meaningful learning task, fifteen logical and numerical series items of the following kind were devised:

(1)	N M L K J	(*a*) G	(*b*) H	(*c*) I	(*d*) J
(2)	1 A Z 2 B Y 3 C	(*a*) Z	(*b*) 4	(*c*) D	(*d*) X

Under Condition A the series on the left was to be completed by selecting one of the four choices on the right and punching the appropriate space on the punchboard. Condition B exposed the learner to two choices, the distractor being selected at random from the three used in A. Under Condition C, *S* was faced with the correct choice only, which was still to be punched in the appropriate space. This type of task was selected on the grounds that it involved seeing relationship such as are found in many school learning problems, particularly in mathematics.

For the rote learning task, *S*s were presented with fifteen rare French words whose correct translation was to be selected from the four, two, or one English words available, depending on whether *S* had been assigned to Condition A, B or C.

The punchboards were cardboard templates containing sixty-four small holes and covered by a paper answer sheet with spaces for responding to sixty-four four-choice questions. A thick piece of blotting paper was inserted between cardboard and answer sheet to prevent the discovery of the correct response by 'feeling' the hole.

Procedure.—*S*s were assigned randomly to the three conditions by the manner in which the question sheets were distributed. Familiarity with the punchboard and the kinds of tasks to be learned was gained by fifteen-minute practice on simple examples. *S*s were instructed to read the question, select an answer, and, on the punchboard, to punch the space which corresponded to that answer with a stylus. If incorrect, *S* was to continue until he found the correct response. *S*s were told to learn the correct answers to all the experimental examples since they would be tested on them subsequently. Directions were suitably modified for the no-choice groups. Since all *S*s faced multiple-choice questions in the practice exercise and at least one of the experimental learning sessions, it was felt that the four-choice questions used in the end-test would be reasonably familiar for *S*s in all conditions.

Five minutes were found adequate for each group to work through and learn each set of tasks, with a one-minute break in between the two sessions. The order of tasks was randomly assigned. When time was called, learning materials were collected and an end-test consisting of the same problems, all in four-choice form, was distributed immediately. These problems differed from those faced by *S*s working under Condition A, in that the order of items and the order of options within each item were rearranged in random fashion. This precaution was taken to reduce the possibility that items might be recalled because of their location or serial position in the punchboard exercises. In the end-test, *S*s merely encircled the option they selected, and hence were given no feedback.

Results.

The mean number of correct responses obtained on the rote learning task from the punchboard exercise and the end-test, for Conditions A, B and C were 6·44, 9·32, 15·00, and 11·95, 12·36, 14·09, respectively. As intended, the four-choice condition produced lower initial scores on the punchboard exercise than did the two-choice condition, while the no-choice group recorded no errors. To test the hypothesis that higher end-test scores would be obtained by the low error groups, three one-tailed t tests were performed. As predicted, Condition C produced significantly higher means than both A and B (t=3·39, 2·93, respectively). The difference between A and B was not signficant (t=0·59).

For the meaningful series task, under Conditions A, B and C, *M*s on the punchboard were 10·82, 12·45 and 15·00; on the end-test *M*s were 11·09, 11·59 and 10·05. In this case the end-test trend was quite different, although the differences between conditions were not significant. Since the direction of difference was not predicted in the end-test a one-way analysis of variance was performed across conditions, yielding a non-significant F value of 1·61 (df=2/63). Clearly, the first hypothesis is supported, that minimal error is associated with better retention in rote but not meaningful learning tasks.

A second hypothesis, that errors, once made, tend to be repeated to a greater extent in rote than in meaningful materials was examined by several item analyses performed on the end-test. The responses of *S*s in Condition B were selected for this analysis, since the attractiveness of their single wrong response to each incorrect item could be evaluated by comparing its frequency of choice with that of the two new distractors.

The first analysis showed that 24·34 per cent. of the rote learning errors made on the punchboard by the twenty-two *S*s in Condition B were repeated by those same *S*s in the end-test. In only 4·00 per cent. of the 152 cases where errors had occurred was a new, but to other groups equally plausible, distractor, chosen. It is thus clear that the wrong response was made more attractive by being selected in the initial learning stage. Since the error repetition rate did not exceed the chance expectation of 25 per cent., it appears that the effect of punching the wrong response before the right was not so much to increase the probability of error as to dictate the pattern of errors in the end-test.

By contrast with these results, an analysis of the fate of the punchboard errors made in the meaningful series problems showed that only 10·00 per cent were repeated while 13·33 per cent. were altered to one of the two new distractors. The previous selection of a wrong response apparently made it no more nor less attractive than the new options. Thus, the Thorndikian tendency to repeat errors did not apply in this kind of learning task.

In a further analysis of the errors made in Condition B tests, it was found that 48·72 per cent. of the errors made on the French end-test were the same as those made on the punchboard, while only 8·69 per cent. of the series end-test errors had been punched before. It was noticeable, too, that the weakest six students in this group accounted for over half the end-test errors, suggesting that these few repetitions may have been the result of an attempt to learn the items by rote.

The effect of both these analyses was to confirm the prediction that errors made in a rote learning task are more often repeated than are errors made in a logical, non-arbitrary task.

III.—EXPERIMENT 2.

A parallel experiment was designed in order to test the same hypotheses with new materials in an elementary school setting. In this experiment, however, retention was tested after a three-day period rather than by immediate retest.

Method.

Subjects.—Ss were eighty-one children from three classes of a city primary school. The age range was from 8 to 11 years, and the median Otis I.Q. was 114.1, ranging from 84 to 145.

Materials.—The meaningful learning task consisted of fifteen verbal analogies typical of those found in intelligence tests, e.g., Bow is to arrow as pistol is to (*a*) bullet, (*b*) holster, (*c*) rifle, (*d*) trigger.

A spelling list of fifteen different words was used for the rote learning task. Since most of the words were not phonetically spelled, it was felt that the learning task was more a matter of making arbitary connections, and therefore, qualified as rote learning, e.g., (*a*) pnumonia, (*b*) pnewmonia, (*c*) pnuemonia, (*d*) pneumonia.

All materials were selected after being presented in four-choice form to a comparable class of 9 and 10-year-old children in order to determine difficulty levels and to isolate poor distractors.

The punchboards used were the same as those employed in Experiment 1.

Procedure.—Instructions and procedures for the learning phase were identical to those of Experiment 1. The end-tests, however, were not administered until three days had elapsed. These tests were designed according to the same principles as those used in the first experiment.

Results.

The mean number of correct responses obtained on the rote learning task from the punchboard exercise and the end-test, for Conditions A, B and C were 5·37, 9·04, 13·96, and 5·11, 4·55, 6·19, respectively.

One-tailed *t* tests performed on the end-test results produced *t* values of 1·69 and 2·04 between the C Condition and A and B, respectively. Both these values reach significance at the ·05 level (df=52) and although not as large as those of Experiment 1, they are consistent with them, and with the hypothesis that low error rates facilitate rote learning tasks.

Means on the meaningful learning task for Conditions A, B and C were: on the punchboard, 6·56, 9·59 and 13·63; on the end-test 9·44, 10·00 and 9·70, respectively. The end-test results for the meaningful task again showed no significant differences in retention between conditions (F=0·19, df=2/24).

The second hypothesis concerning the tendency of pupils to repeat their errors was examined by item analysing the responses of the two-choice group.

Of the 166 errors made initially in the learning of the rote material, fifty-six (or 33·75 per cent.) were found to be repeated. This figure is significantly more than the chance reference level of 25·00 per cent. ($CR=2{\cdot}39$, $p<{\cdot}01$).

In the meaningful analogies task, 144 errors were made initially on the punchboard. Of these forty-four (or 30·56 per cent.) were repeated in the end-test. Although this is considerably more than in the previous experiment, it is not significantly greater than the chance expectation level ($CR=1{\cdot}45$). Closer analysis showed that 60 per cent. of the errors repeated were produced by eight pupils whose initial and final scores on the punchboard were low enough to suggest that they were learning by rote. The initial and final mean scores of these eight pupils were 7·88 and 7·38, respectively, as compared with 10·31 and 11·11 gained by the remaining nineteen pupils in the two-choice group.

IV.—CONCLUSIONS.

The findings of both these experiments support the general hypothesis, discovered by Thorndike and emphasised by Skinnerian programmers, that error are undesirable in rote learning, since they tend to be repeated, even with immediate feedback. With logically related, meaningful problems, however, the incidence of errors was found to be unrelated to ultimate performance. Error repetition was virtually absent in training college students and was apparent only among the poorer elementary school children who appeared to be learning without understanding.

Since these findings were made with two kinds of rote and meaningful materials in each case, and with two widely different age groups, the conclusions seem to have general significance for programmers and teachers. Error rates as a criterion for evaluating logical and meaningful programmes are probably irrelevant. The significant variable is performance in an end-test, which is largely independent of error rates within the programme. Moreover, the advantage gained by freeing the programmer from the restricting assumption that errors are to be kept to an absolute minimum should make for more imaginative programmes and more interested students. The ' pall effect ' of working through long, easy, repetitious programmes is now too well documented for programmers to overlook this advantage. The insertion of a few challenging questions which provoke error, may actually help rather than hinder the student.

On a more general level, teachers, who, after Thorndike, have been exhorted to ' start right ' and ' avoid errors ' in learning, might be further advised to restrict such policies to rote learning. Thorndike's tendency to play down ' discovery learning ' on the grounds that misconceptions acquired in the process are difficult to eradicate, finds little support beyond rote learning in the present study.

V.—REFERENCES.

AUSUBEL, D. P. (1963). *The Psychology of Meaningful Verbal Learning.* New York: Grune and Stratton.

CROWDER, N. A. (1964), On the differences between linear and intrinsic programming. In J. P. de Cecco (Ed.) *Educational Technology.* New York: Holt, Rinehart and Winston., pp. 142-151.

HILGARD, E. R. (1956). *Theories of Learning.* New York: Appleton-Century-Crofts.

KAESS, W., and ZEAMAN, D. (1960). Positive and negative knowledge of results on a Pressey-type punchboard. *J. Exp. Psychol.*, 60, 12-17.

SKINNER, B. F. (1958). Teaching machines. *Science*, 128, 969-977.

THORNDIKE, E. L. (1932). *Fundamentals of Learning.* New York: Teacher's College, Columbia Univer.

WILLIAMS, J. D. (1963). Aspects of programmed instruction: I—Some notions and arguments. *Educ. Res.*, 5, 163-178.

(*Manuscript received 2nd September*, 1965)

A TRAINING PROCEDURE FOR ACQUISITION OF PIAGET'S CONSERVATION OF QUANTITY: A PILOT STUDY AND ITS REPLICATION*

By IRVING E. SIGEL †
(*Merrill-Palmer Institute*)

ANNEMARIE ROEPER
(*City and Country School*)

AND FRANK H. HOOPER
(*Center for the Study of Cognitive Processes, Wayne State University*)

SUMMARY. This study represents an initial investigation of the relationship between logical, pre-requisite operations and Piaget's conservation tasks. The hypothesis considered special training on multiple classification, multiple relations, and reversibility, as a means of inducing conservation acquisition or improvement. Twenty nursery school children between the ages of 4·3 and 5 with I.Qs. over 130 were assigned to training and control groups. Ten of the subjects made up a pilot study and ten constituted a replication sample. All subjects were given four conservation of continuous quantity tasks, i.e., substance, liquid substance, weight, and volume. The training subjects received classification, relationality, and reversibility enrichment experiences in a small group setting. The control subjects had non-relevant, interpolated activity.

Post-testing shows clear differences between the training and control subjects. Considering the subjects who indicated a complete lack of pre-test conservation ability, seven of eight Training Group children improved in later performance. No change was shown for seven Control subjects. Significant post-test differences, were found for substance and weight conservation. In addition, the training subjects' performances indicated a greater awareness of relevant attributes and increased verbal sophistication within the test situation.

The failure of previous conservation training attempts is attributed to a lack of concern for the stage-related, pre-requisite operations. Direct training on Piagetian tasks may be unnecessary if attention is directed to the logical precursors of specific levels of cognitive development.

I.—INTRODUCTION.

CONSERVATION is defined by Flavell (1963, p. 245) as " the cognition that certain properties (quantity, number, length, etc.) remain invariant (are conserved) in the face of certain transformations (displacing objects or object parts in space, sectioning an object into pieces, changing shape, etc.)." This process is considered by Piaget as " a necessary condition for all rational activity " (Piaget, 1952, p. 3). On the basis of extensive studies, Piaget identifies three stages of development for each type of conservation: stage 1, in which no conservation is found, and child focuses in this instance on irrelevancies; stage 2, a transitional period, in which the child is dominated by perceptual appearances and in which conservation may or may not appear; stage 3, the stage of natural conservation, in which the child readily and logically demonstrates understanding of invariant properties in the face of transformations. The stages are presumed to be fixed and invariant, irrespective of the properties under study. There is an additional developmental sequence which exists across the different quantity subclasses. Conservation of substance appears at about 8 to 10 years of age, preceding conservation of weight which is apparent between the ages of 10 to 12. Functional acquisition of volume occurs after the age of 12. The invariant sequence of the development of substance, weight and volume conservation has generally been verified, with certain qualifications, in a number of studies by Elkind (1961a, 1961b, 1961c), Kooistra (1963), Lovell (1961) and Smedslund (1961a, 1961b, 1961c, 1961d, 1961e, 1961f).

---

* A partial report of this study was read at the Conference on " New Media in Education, Business and Industry," Wayne State University, Detroit, Michigan, January, 1965.

† 71, East Ferry Avenue, Detroit, Michigan, 48202.

Can conservation be induced in the child? Piaget (1952) holds that for conservation to appear the child must be able to perform the following operations: multiple classification, multiple relationality, atomism, reversibility, and seriation. We contend that conservation can be induced if the training procedures embody these prerequisite operations.

II.—PREVIOUS TRAINING RESEARCH.

The majority of previous attempts at conservation modification or acceleration have dealt with areas other than quantity concepts. These include number conservation, Wohlwill (1960), Wohlwill and Lowe (1962) and area conservation, Beilin and Franklin (1962). Training results in these studies have essentially been negative. Piagetian concept acquisition has not proved amenable to laboratory training techniques based upon traditional learning approaches, e.g., see Flavell (1963, pp. 370-379). The studies which have attempted to induce conservation generally, provide no information as to the child's ability to employ any or all of the prerequisite operations. They merely indicate the extent to which the child can conserve (Bruner, 1964, and Smedslund, 1961a). Further, these investigators shed no light on the relationship between the mental operations, e.g., multiple classification and conservation.

Piaget assigns a clearly subordinate role to language as a necessary condition in the development of conservation, holding that changes in cognitive structure are not accomplished via verbal enrichment or sophistication. Bruner, however, contends that acquisition of conservation is related to linguistic experiences. He reported a study by Frank who tested the Bruner hypothesis that " improvement in language should aid this type (conservation) of problem solving " (Bruner, 1964, p. 5). Activation of language would enable the child to be less dominated by perceptual forces in the setting, less inhibited in utilizing symbolic processes, and consequently, be able to deal with the conservation problem. Frank, using subjects age 4 to 7, tested the children on the classic paradigm of conservation of continuous quantity (liquid), in one case with the materials in full view, and for several conditions in which the beakers are screened. In the latter condition the child is asked to verbalize. Increases in conservation responses occurred under the screened conditions and the older subjects maintained correct responses in post-test and transfer task situations. These results are offered by Bruner as support for the significance of verbalization in conservation acquisition (Bruner, 1964, pp. 6-7).

Unfortunately, the report of this experiment is condensed and open to alternative interpretations. Selective attention toward the salient cues of constant water quantities, especially in the screened situation, may be the operative factor, irrespective of any verbalization that occurs. Practice effects via the recall of the treatment condition would not necessarily demand a language activation rationale. In particular, we would have to know how the experimenter examined the influence of problem-specific queries across a number of conservation settings. Although Bruner reports no significant improvements for pre-test—post-test controls, Smedslund has found changes following task administration, irrespective of training methods employed. The possibility of learning within the task situation still remains.

Another set of studies involving the acquisition of the conservation of weight has been reported by Smedslund (1961a, 1961b, 1961c, 1961d, 1961e, 1961f). Working with children from 5 to 7, Smedslund attempted to train children to conserve weight by employing two training procedures. In one situation he reinforced trials in the conservation of matter. He altered the form of one of the balls of clay and asked the child to predict which one would weigh more. The child's prediction was tested directly by placing the object on a scale in the child's presence. A second procedure was to add or subtract pieces of plasticine and show the child the relative effect of these actions. The assumption in this latter approach was that if the addition-subtraction schema were exercised, the child would more easily acquire conservation of weight. The results for this study were essentially negative.

One experiment in this series by Smedslund is of considerable interest in that it provides exciting possibilities for inducing conservation by the creation of cognitive conflict, i.e., competing cognitive systems. Smedslund reasoned that creation of such conflict would induce cognitive reorganization. Two kinds of transformations were employed, i.e., deformation by changing the shape, and addition and subtraction of quantity. Thus, if the child was inclined to think that flattening out a piece of clay increased its size, and that subtracting a piece of it would decrease its size, the experimenter would employ both actions at the same time and ask the child whether the transformed and subtracted item was the same or less than the standard. Presentation of two conflicting approaches is presumed to give the subject a situation in which he has to pause and decide between the conflicting possibilities. Of all the conditions Smedslund employed, to induce conservation, the conflict paradigm was the most successful.

A recent study by Gruen (1964) compared verbal pre-training on number conservation, in combination with direct reinforced practice, against a verbal cognitive conflict paradigm drawn from Smedslund's approach. Although training effects were generally small, a significant difference between verbal conflict training and control groups was demonstrated. In addition, Beilin (in press) found a verbal-rule instruction method to be effective in the inducement of length and number conservation. Tests on an area generalization task revealed a notable absence of any transfer effects.

It may be that facilitating as cognitive conflict is, it may have to be more intensive as well as extensive. In Smedslund's studies only five of thirteen subjects manifested changes in conservation ability. Why not all, in view of their common exposure to the training? Our contention is that conservation can be induced only by training the child in the prerequisites to such mental operations. Such training may, in fact, induce conflict since it exposes the child to experiences which force him to reorganize cognitively.

The potential role of operational reversibility in mediating conservation development is indicated by Eifermann and Etzion (1964) for adult subjects, and by Wallach and Sprott (1964) for young children. The latter study utilized a reversibility training approach very similar to that found in the final sessions of the present enrichment progam. Experimental training groups of 6-7-year-old children were found to differ significantly from controls on discontinuous quantity conservation and numerical correspondence responses.

III.—RATIONALE FOR PROPOSED TRAINING PROGRAM.

To conserve, the child must be able to emancipate himself from the perceptual demands of the situation, to know that the size or shape of the items can be changed, i.e., be transformed, and still not lose every one of the initial attributes, and to acknowledge that objects are multidimensional. Given that the child learns that objects can be classified and reclassified on the basis of any one of the single attributes, he is now in a position to learn that objects can be classified and reclassified on the basis of two or more attributes. Combining two or more dimensions of objects signifies the beginning of multiplicative relations, an operation that leads to the creation of a large array of classes. Since objects have a large number of possible class memberships, classifications are, to a degree, relative. The creation of classes becomes dependent on the criteria selected as bases for classification. An interesting by-product of this is the awareness that class membership of an item may be defined by the individual, since it is he who decides what attributes to use.

The above operations are presumed to be critical for the acquisition of conservation. What about the schemas? Piaget proposes the schema of atomism, the awareness that objects are made up of minute particles. The transformations are said to relocate these particles in space without changing their number. It is our contention that such knowledge is unnecessary for the level of comprehension of conservation under discussion here.

In summary, then, we conclude that conservation is possible only when the mental operations of ***multiple classification***, ***multiple relationality***, and ***reversibility*** are present as the cognitive structure basic to adequate conservation performance. The goal of our program was to train children in these operations.

IV.—PLAN OF THE INVESTIGATION.

(*a*) *Subjects.*

Ten children ranging in age from 4·9 to 5, were divided in two groups of five and comprise the Pilot sample. One group was designated Training Group and the other Control. No significant age differences were found between the two groups. The mean Stanford-Binet I.Qs. were 149 for the T group and 152 for the C group. Unfortunately, school withdrawals prevented post-testing of one training and two control subjects. All children were in regular attendance at a nursery school at which the second author was Director. This school provides a rich and stimulating curriculum.

(*b*) *Conservation Tasks.*

Each of the ten children was given conservation tasks of continuous quantity (Plasticine) for substance, weight, and volume, and one conservation of liquid substance. The tests were administered in the classical way. A child passed a conservation task when he could indicate that at least in two of the three transformations the two items were still identical in terms of the property under discussion. Post-tests were administered two weeks after the training period.

(*c*) *Teaching Procedures.*

Of particular significance in this report is the teaching procedure used. The Training Group gathered in Mrs. Roeper's office for their sessions, each of which was tape recorded. The children were told that they were going to do some work, and nothing was said about the relationship between the work in Mrs. Roeper's room and the testing experiences.

On the basis of the theoretical position presented earlier, the training focused on multiple labelling which was followed by multiple classification and multiplicative relations. Finally, reversibility was introduced. In Table 1 we have a verbatim account of the teacher introducing an object, in this case a banana, and the encouragement toward labelling. This is followed by another object, an orange. This same procedure is followed. Having focused on the characteristics of each, the teacher presents two similar objects and asks the children to label differences between the tangerine and the orange. Having identified the difference, she now proceeds to discuss similarities.

TABLE 1

PORTION OF VERBATIM TRANSCRIPT OF A TRAINING SESSION DEALING WITH MULTIPLE ATTRIBUTES OF OBJECTS.

Teacher : Can you tell me what this is, Mary ?
Mary : A banana.
Teacher : What else can you tell me about it ?
Mary : It's straight.
Teacher : It's straight. What else ?
Mary : It has a peel.
Teacher : It has a peel. . . . Tom, what can you tell me about it ?
Tom : Ummm . . . It has some dark lines on it.
Teacher : Uh-hugh.
Tom : It has some green on it.
Teacher : What can you do with it ?
Tom : You can eat it !
Teacher : That's right ! . . . Now let's see . . .
Children : . . . I love bananas !
Teacher : What is this ?
Children : An orange.
Teacher : Is it really an orange ?
Children : Uh-huh. . . . Yes.
Teacher : Look at it closely.
Child : It's an artificial one.
Teacher : Oh, that's right, it's an artificial one . . But, what else can you tell me about it ?
Children : You can eat it . . It is round . . .
Teacher : Uh-huh.
Children : . . . Orange.
Teacher : That's right !
Child : It has a stem.
Teacher : Now, look at this one . . . What's this ?
Children : An orange . . orange.
Teacher : And what can you do with it ?
Children : You can eat it . . . and it's round . . .
Teacher: It is round . . .
Child : It has a peel . . .
Teacher : It has a peel . . . Now, look at these two things. Are they the same ?
Children : No.
Teacher : What's different ?
Children : This one . . . this one here is pressed in on the side a little . . . this one is lighter.
Teacher : Do you know what this really is ? This is a tangerine . . . and this is an orange. Now tell me in what ways they are alike.
Children : This is smaller and that's bigger.
Teacher : I said, " In what way are they alike ? "
Children : They are both round . . . they both have a stem . . . both orange.
Teacher : They both have a stem, both round, both orange : Anything else alike about them ?
Child : They're both fat.
Teacher : Uh-huh. What can you do with them ?
Children : We can eat them . . .
Teacher : We can eat them . . . Now, tell me, what's the same about all these things ?
Child : These are round, but this isn't.

Teacher : I said, what is the same about them, not what's different about them.
Children : They're both round . . . they're round . . . they're round . . . and they are both artificial.
Teacher : They're all artificial, and, . . . are they all round ?
Child : No.
Teacher : What about the banana. ?
Child : It's straight.
Teacher : But, . . . tell me something else that's the same about all of these things.
Child : . . . They have . . . all have a peel.
Teacher : That's right, too, but what can you do with all of them ?
Children : You can eat them !
Teacher : That's right ! That's the same about every one of them. Do you have a name for all of them ?
Children : Yes !
Teacher : What ?
Child : A banana.
Teacher : A banana ? No, . . . is there something that you can call all of them ?
Children : Fruit . . . fruit.
Teacher : And what's the same about all fruit ?
Children : They are all round except bananas.
Teacher : No, . . . why do you call all of these things fruit ?
Children : Because you can eat them.
Teacher : You can eat them.
Children : And they are food.
Teacher : And they are food. If I had a piece of bread here, would that be fruit too ?
Children : No.
Teacher : Why not ?
Children : Because it is not sweet . . . not round . . .
Teacher : Because it is not sweet. I think that's a good reason . . . and, you eat bread too ?
Children : Yes.
Teacher : But it still not a fruit . . . right ?
Children : Yes.
Teacher : Now, can you tell me again what this is ? We talked about it yesterday.
Child : A pencil.
Teacher : What else can you tell me about it ?
Children : It's round. . . . You said you were going to put it in . . .
Teacher : That's right . . . Ah . . . Tom, what is this ?
Tom : Chalk.
Teacher : What else can you tell me about it ?
Tom : It's white.
Teacher : Gail, tell me, what's the same about these two things ?
Gail : They're both round.
Teacher : What else ?
Gail : . . . . . . Ummm . . .
Teacher : John, tell me, what's the same about these two things ?
John : . . . both write.
Teacher : That's right ! There are two things that are the same about it. Tell me what they are ?
John : Well. . . . I don't know . . .
Teacher : What are they, Mary ?
Mary : They're round and they write.
Teacher : Very good !

We assume that directing the children's attention to the similarities, as well as the differences of objects, facilitates perception of comparison and distinction. The children begin to ennumerate the comparison and each one is accepted by the teacher as legitimate. This is contrary to some of their experiences on intelligence tests, for example, where such acceptance is not possible or responses are not given the same reward. Having ascertained and verbalized apparent similarities such as shape and color, the teacher introduces another basis of classification, in this particular case, functional attributes. Having worked with the two different objects, yet of the same general class, she now introduces the original banana, which is different, and focuses on similaritie again. The process of naming the additional attributes, searching for class labels, and defining criteria of class membership goes on until the group, under the teacher's guidance and stimulation, has exhausted the possibilities. Following this, the children were introduced to the idea of multiplicative relations in which two criteria are co-ordinated as a basis of classification, e.g., " Can you think of two things that you are at the same time ? " Starting with the personal, that is the ego-centric, the teacher moved to non-personal objects.

Reversibility was the final operation to be introduced. The children were given a constant number of pennies, which were divided amongst them. The goal here is to demonstrate that irrespective of how the pennies are divided or arranged on the table, the total remains the same and can be recombined to form a pile equal to the original one.

The opportunity to use language and language formulas, and to provide schemas in the context of learning particular information was deliberate. We also believe that opportunities for cognitive reorganization were provided, in that the children learn two or more attributes of an object which can be combined to define another class. Thus, through structural teaching involving verbalization and demonstrations, we sought to bring to the children's attention the fact that objects have multiple characteristics (multiple classification), that these can be combined in various ways to produce new categories (multiple relations), and that categories of objects can be reorganized and brought back to the original (reversibility).

Although each of these sessions was planned in advance, a rigid schedule could not be followed. The children did not always conform to the teacher's expectation, at times they gained insight so quickly that the teacher had to vary her procedure. The sessions lasted only twenty to thirty minutes and were discontinued as soon as the teacher noticed that the children seemed to lose interest. The Control Group would also come into Mrs. Roeper's office, but were engaged in discussions of social studies problems. They would discuss such matters as community helpers and their social roles. In this way, both groups were given special treatment, differing in specific content. It will be noticed that the training procedure did not focus directly on the tasks that were used in the post-testing, nor was any reference made to that situation. We believe, as far as the children were concerned, they were unaware of the relationship between the two.

V.—INITIAL RESULTS.

The results of the pilot sample are shown in Table 2. Each child in the Training Group was able to solve correctly at least one more conservation task in the post-testing than in the pre-testing session. The most dramatic change is that of Gail who responded correctly to the post-test problems involving substance, weight, and volume. Tom and Mary showed smaller gains, but nevertheless, did do better on the post-training testing than on their pre-training tests.

For the Control Group children, two out of the three were unable to solve any of the conservation tasks on the pre-training tests and were also unable to solve any of the tasks on the post-training testing. One child in the Control Group, Em, who was able to solve two of the conservation tasks, was able to solve all of them in the post-training.

In discussing the findings, we must keep in mind that the children are of comparable I.Q. chronological age, social status and educational level. The answers to the differences in the abilities to resolve the conservation tasks cannot be attributed to any of these variables.

VI.—RESULTS OF REPLICATION STUDY.

The results of the first training session were sufficiently encouraging to warrant a replication. In the selection of the second sample, we made certain that none of the ten children could conserve. The children were assigned randomly to the training and control groups, the average ages were 4·3 and 4·5, respectively, and mean Stanford-Binet I.Q. scores were identical, 143·8. Identical testing and training procedures, as described in the first situation, were used.

The results of the pre- and post-conservation testing are presented in Table 2.

TABLE 2

RESPONSES OF TRAINING AND CONTROL GROUPS IN TWO CONSERVATION EXPERIMENTS.

STUDY I								
Training Group								
	Pre-Testing				Post-Testing			
Subjects :	Sub.	Liq. Sub.	Wt.	Vol.	Sub.	Liq. Sub.	Wt.	Vol.
Gail	0	0	0	0	+	0	+	+
Tom	0	0	0	0	0	+	+	0
Martha	+	0	0	0	+	0	+	0
Mary	0	0	0	0	+	0	0	0
Control Group								
Ma	0	0	0	0	0	0	0	0
Je	0	0	0	0	0	0	0	0
Em	+	†	0	0	†	†	+	+
STUDY II								
Training Group								
Ruth	0	0	0	0	+	+	†	0
Joby	0	0	0	0	0	0	0	0
Nelson	0	0	0	0	+	0	0	0
Jody	0	0	0	0	+	+	+	0
Tracy	0	0	0	0	0	0	+	0
Control Group								
Al	0	0	0	0	0	0	0	0
Su	0	0	0	0	0	0	0	0
Car	0	0	0	0	0	0	0	0
Ja	0	0	0	0	0	0	0	0
Ji	0	0	0	0	0	0	0	0

In the pre-test, not one child was capable of conserving. In the post-test, we find that four of the five children were able at least to conserve one or more properties. It is particularly interesting to note that Child 5 was able to conserve weight but not substance or liquid substance. Beyond the objective pass-fail considerations, the training subjects revealed a heightened awareness and verbal sophistication within the post-test setting. Thus, the children not only increased in their ability to conserve correctly, but also showed an increase in their ability to verbalize the underlying operations and the salient dimensions of the criterion test situation. The Control Group showed no change in conservation ability or verbalization skills.

VII.—DISCUSSION AND CONCLUSION.

How can we explain these particular findings? The only significant difference in the educational experience of the Training and Control Groups was the training session during the five-week period. These findings indicate that given training in such operations as multiple classification and reversibility, conservation, at some level, has a greater probability of appearing.

We do have the single case of the child in the first Control Group who showed a change over the course of the five-week period. This may be due to the child's already exhibited ability to conserve which was highlighted by the experimental situation. Yet, we find no such phenomenon in the second Control Group, e.g., no conservation in the pre-test, none in the post-test. This strongly supports the contention that the child who can conserve is in the process of demonstrating it in more situations. If we discount, then, the child Martha in the Training Group, and Em in the Control Group, who had shown pre-test conservation, we find that the children in the first Training Group who had no conservation did acquire conservation ability, whereas this was not the case in the Control Group. The children who show no such ability are able to solve conservation problems *only* after the training experiences. This suggests that the testing situation may not be an influence unless some ability to profit from the experience is present. Combining the two samples, eight training and seven control subjects indicated pre-test non-conservation. There were no post-test successes for any of the seven control subjects. In contrast, of the eight training subjects, five conserved substance, three conserved liquid substance, five conserved weight, and one subject succeeded on the volume task. Pass-fail contingency comparisons show differences between training and control performances on the substance and weight cases, e.g., Fisher Probability Test significance level of ·025.

The sequence of development, described by Piaget and confirmed by Elkind, Kooistra, Lovell, and Smedslund, appears in this study with the exception of one subject, Tracy, in our replication sample. The degree to which the enrichment experiences were assimilated by the Training Group is reflected in the type of response given to the inquiry questions in the post-training session. The children verbalized their explanations in an articulate way, employing statements of reversibility, for example, as explanations. During the post-testing session, in response to the question, " Do I have more or less, or the same amount of clay? " the answer was much more complex than in the pre-training session. The child answered, " This piece is the same as this piece (points to the shape-transformed stimuli), it looks the same, it is the same. It is the same clay, it is turned into a ball just like that (points to the comparison stimuli), then you changed it into something else. It looks the same, the same amount of clay." What is of interest in this case is that the explanations are at levels expected as a consequence of training experience. The children who did not have any training showed little change in verbalization patterns from initial to post-training test situations. Even when giving the wrong answer, the Training Group was more articulate and fluent than the Control Group.

The results demonstrated in the ability to articulate as well as in the ability to answer the questions correctly are to be explained primarily as a function of the training experience. Such a conclusion is warranted on theoretical as well as empirical grounds. Reading the transcripts, it becomes readily apparent that the children become sensitized to the multiplicity of attributes of objects and to the possibility of co-ordinating combinations of attributes. They seem to be encouraged by such an experience to make inferences, that is, to view objects in terms of their classificatory relationships to other objects. The encouragement to verbalize their ideas about objects and to direct attention to the multidimensionality of objects may have played an important part in the child's assimilation of this new knowledge. Such encouragement to verbalize, in conjunction with the prerequisite operations, may combine to provide the appropriate experiences necessary for the child to acquire conservation. In this sense, Bruner's emphasis upon the verbal aspect of the situation may well be correct.

The quantitative and qualitive analyses of the results of this study provide support for the basic hypothesis that training programs focusing on prerequisites for relevant cognitive operations influence the resultant cognitive structures. This modification of cognitive structure, directed by such a training procedure, provides for the emergence of new sets of abilities and demonstrates the interdependence of cognitive stages and precursors. On the basis of this position, conservation *per se* need not be taught directly. Possessing the necessary prerequisites, children may ' discover ' conservation as a principle and apply it to substance, weight and volume. Increased affirmation of their competence in a specific area could be obtained through extended practice. They already have the necessary logical structure, in this instance, to master conservation of continuous quantity. Children having experience with conservation of continuous quanity could then be exposed to discontinuous quantities, number, etc. In this way, progress toward comprehensive logical thought can be developed. Insofar as stage placement within a given quantity conservation type is concerned, the present training subjects were probably at a transitionary or intuitive level. It appears reasonable that transitionary stage individuals would derive the greatest benefit from the training experiences. Unfortunately, no attempt was made to determine stage location in this study.

VIII.—REFERENCES.

BEILIN, H. (in press). Learning and operational convergence in logical thought development. *J. Exp. Child Psychol*

BEILIN, H., and FRANKLIN, I. C. (1962). Logical operations in area and length measurement, age and training effects *Child Developm.*, 33, 607-618.

BRUNER, J. S. (1964). The course of cognitive growth. *Amer. Psychologist*, 19, 1-15.

EIFERMANN, R. R. and ETZION, D. (1964). Awareness of reversibility: its effect on performance of converse arithmetical operations. *Brit. J. Educ. Psychol.*, 34, Pt. 2, 151-157.

ELKIND, D. (1961a). The development of quantitative thinking : a systematic replication of Piaget's studies. *J. Genet. Psychol.*, 98, 37-46.

ELKIND, D. (1961b). Children's discovery of the conservation of mass, weight, and volume : Piaget's replication study II. *J. Genet. Psychol.*, 98, 219-227.

ELKIND, D. (1961c). Quantity conceptions in junior and senior high school students. *Child Developm.*, 32, 551-560.

FLAVELL, J. H. (1963). *Developmental Psychology of Jean Piaget*. N.J. : Van Nostrand.

GRUEN, G. E. (1964). *Experience Affecting the Development of Number Conservation in Children*. Unpublished doctoral dissertation, Univer. of Illinois.

HUNT, J. McV. (1961). *Intelligence and Experience*. New York : Ronald Press.

KOOISTRA, W. (1963). *Developmental Trends in the Attainment of Conservation, Transitivity, and Relativism in the Thinking of Children : A Replication and Extension of Piaget's Ontogenetic Formulations.* Unpublished doctoral dissertation, Wayne State Univer.

LOVELL, K. (1961). *The Growth of Basic Mathematical and Scientific Concepts in Children.* New York : Philosophical Library.

PIAGET, J. (1952, 3rd impression, 1964). *The Child's Conception of Number.* London : Routledge and Kegan Paul.

SMEDSLUND, J. (1961a). The acquisition of conservation of substance and weight in children : I.—Introduction. *Scand. J. Psychol.*, 2, 11-20.

SMEDSLUND, J. (1961b). The acquisition of conservation of substance and weight in children : II.—External reinforcement of conservation of weight and of the operations of addition and subtraction. *Scand J. Psychol.*, 2, 71-84.

SMEDSLUND, J. (1961c). The acquisition of conservation of substance and weight in children : III.—Extinction of conservation of weight acquired ' normally ' and by means of empirical controls on a balance. *Scand. J. Psychol.*, 2, 85-87.

SMEDSLUND, J. (1961d). The acquisition of conservation of substance and weight in children : IV.—Attempt at extinction of the visual components of weight concept. *Scand. J. Psychol.*, 2, 153-155.

SMEDSLUND, J. (1961e). The acquisition of conservation of substance and weight in children. V.—Practice in conflict situations without external reinforcement. *Scand. J. Psychol.*, 2, 156-160.

SMEDSLUND, J. (1961f). The acquisition of conservation of substance and weight in children : VI.—Practice on continuous *versus* discontinuous material in conflict situations without external reinforcement. *Scand. J. Psychol.*, 2, 203-210.

WALLACH, LISE and SPROTT, R. L. (1964). Inducing number conservation in children. *Child Developm.*, 35, 1057-1071.

WOHLWILL, J. F. (1960). A study of the development of the number concept by scalogram analysis. *J. Genet. Psychol.*, 97, 345-377.

WOHLWILL, J. F. and LOWE, R. C. (1962)., Experimental analysis of the development of the conservation of numbers. *Child Developm.*, 33, 153-157.

(Manuscript received 26th August, 1965)

LEVEL, SPEED AND PERSONALITY FACTORS IN THE INTELLECTUAL PERFORMANCE OF YOUNG CHILDREN

BY AUBREY J. YATES*
(*University of New England*)

SUMMARY. Part of Progressive Matrices (1947) and an Arithmetic test were administered with initial and final time limits to a group of 100 9-year-old pupils. It was shown that, as in previous studies, with university students, groups of subjects could be identified who were slow but accurate workers whose intellectual level was severely underestimated by imposed time limits. It was also shown that the same subjects tended to be handicapped by imposed time limits on both tests. No relationship between preferred work-method and neuroticism or extraversion scores on the Junior M.P.I. was found.

I.—INTRODUCTION.

IN previous studies (Yates, 1963, 1966) it was shown that a small, but possibly important, group of university students was severely handicapped in performance on Progressive Matrices (1947) by the imposition of an arbitrary time-limit, as shown by a marked increase in score when additional time was allowed. It was further shown that this group of students was not inferior in *level* of intelligence but rather worked slowly but accurately, since on the Nufferno Level test they completed fewer items while obtaining a level score equal to that of faster-working students. In fact, performance on Progressive Matrices in unlimited time suggested that many of these students were of above-average intelligence in relation to university norms for the test. It was also shown that these students tended to be handicapped when a time limit was imposed on an objective mathematical test.

The groups investigated in the previous studies constituted, of course, a sample which was already highly selected and it was, therefore, feasible to speculate that other students showing the same pattern of slow but accurate work might well be eliminated at selection stages earlier than university level. The possibility that these preferred patterns of work were related to personality factors was also discussed, though no direct evidence on the matter was available from these studies.

The present study was, therefore, intended to investigate whether similar results would be obtained when school children were examined and further to determine whether personality factors were involved.

II.—SUBJECTS AND PROCEDURE.

The subjects consisted of 100 Grade IV (9-year-old) pupils at Mount Pleasant Primary School in Perth. The tests used were Progressive Matrices (1947) and a specially constructed Arithmetic test, suitable for this age level. In the Progressive Matrices test, items 3-10 of the practice book (Set I) and items 1-20 of the test proper (Set II) were used.

The procedure used was as follows : For Progressive Matrices, the subjects were first shown how to work the first three items in Set I. They were then given ten minutes to work items 3-10 of Set I, proceeding immediately to items 1-20 of Set II. At the end of this period they circled the answer last completed and were then given a further ten minutes to complete the test. In the Arithmetic test they were given fifteen minutes, circled the last completed answer, and were then given a further ten minutes.

---

* This study was carried out as an undergraduate student investigation by Noel A. Trotter, under the direction of the author at the University of Western Australia.

On a separate occasion, the Junior Maudsley Personality Inventory (Furneaux and Gibson, 1961) was administered.

III.—HYPOTHESES.

The following hypotheses were to be tested :

(1) That, on the Progressive Matrices, a group of slow but accurate subjects would be identified, who would benefit markedly by the provision of additional time.

(2) That, on the Arithmetic test, a similar group would be identifiable.

(3) That these two groups would contain many of the same subjects.

No predictions were made concerning the personality structure of this group of students, since there is evidence in the literature implicating both neuroticism and introversion-extraversion as possible relevant personality factors (Lynn and Gordon, 1961).

IV.—RESULTS.

(1) *Progressive Matrices.*

Table 1 shows the relationship between score increase in additional time and pattern of scores (passed, failed, not attempted) in the first period for Progressive Matrices. Groups (0, 1, 2), comprising about 60 per cent. of the total sample complete many more items than the remaining groups and show little improvement in score in the additional time. They also tend to obtain rather poor scores. The remaining groups tend to fall into two categories : those groups (7, 8, 13) which also obtain relatively low scores in the initial time, but fail many fewer items than groups (0, 1, 2) and obtain a large increase in score in additional time ; and those groups (3, 4, 5, 6) which obtain the highest scores in the initial time, but also fail many fewer items than groups (0, 1, 2) and obtain a large increase in score in additional time. (Group 10 does not seem quite to fit any of these categories.)

TABLE 1

RELATIONSHIP BETWEEN SCORE INCREASE IN ADDITIONAL TIME AND SCORE PATTERN IN FIRST PERIOD ON PROGRESSIVE MATRICES.

Score Increase ("Group")	Score Pattern at end of First Period			
	Pass	Fail	Not Attempted	N
0	4·7	17·7	7·6	22
1	5·0	12·2	12·8	20
2	5·1	12·7	12·2	19
3	6·8	7·2	16·0	5
4	7·0	7·4	15·6	5
5	7·5	7·5	15·0	6
6	7·8	7·0	15·2	5
7	5·2	5·0	19·8	5
8	5·9	4·7	19·4	7
10	8·6	4·0	17·4	5
13	6·0	4·0	20·0	1

Thus, the results confirm and possibly extend somewhat the previous finding with university students that a sizeable proportion of intelligent students do not reveal their true level when arbitrary time limits are imposed.

(2) *Arithmetic Test.*

Table 2 shows the same relationships for the Arithmetic test. The pattern of results obtained has similarities and differences to those obtained with Progressive Matrices. Since the maximum score obtained by any subject at the end of the first period was eight correct answers, the ' pass ' mark was arbitrarily set at four correct answers for the first period. Since the maximum score obtained by any subject at the end of the second period was sixteen correct answers, the ' pass ' mark, was arbitrarily set at eight correct answers for the total period. The pattern of results then falls into three fairly clear-cut groups. Groups (0, 1, 2, 3, 4) ' failed ' the test in both initial and total periods ; group (5) ' failed ' the test in the first period, but ' passed ' it when additional time was given ; and groups (6, 7, 8, 9, 10) ' passed ' the test in both initial and final periods. The performance of groups (5–10) was, however, clearly underestimated to a considerable degree by the imposition of an arbitrary time limit. The pattern of results differs from that obtained with Progressive Matrices, however, in that the ' handicapped ' groups were markedly superior to the other groups within the initial time limit ; and this superiority was increased bv the additional time limit.

TABLE 2

RELATIONSHIP BETWEEN SCORE INCREASE IN ADDITIONAL TIME AND SCORE PATTERN IN FIRST PERIOD ON ARITHMETIC TEST.

Score Increase ("Group")	Score Pattern at End of First Period			
	Pass	Fail	Not Attempted	N
0	1·2	1·8	15·0	10
1	1·8	2·2	14·0	6
2	1·8	2 0	14·2	5
3	2·7	1·4	13·9	21
4	3·4	1·6	13·0	14
5	3·7	1·0	13·3	24
6	4·5	0·9	12·6	8
7	4·0	1·7	12·3	7
8	6·5	1·0	10·5	2
9	7·0	0·5	10·5	2
10	5·0	0·0	13·0	1

(3) *Relationship between Progressive Matrices and Arithmetic Pattern of Performance.*

The most interesting and important question, however, concerns the hypothesis that the subjects showing a significant increase in performance on the Matrices in additional time will be the same subjects who benefit from additional time on the Arithmetic test. In order to examine this question, the groups shown in Tables 1 and 2 were arbitrarily divided into a low increase

Matrices group and two high increase Matrices groups and then classified into ' pass ' or ' fail ' groups on the Arithmetic test for the first and second periods separately, with results shown in Table 3. The most important result of this analysis relates to groups (7, 8, 13) which were severely handicapped by the time limit imposed on the Matrices. Seventyseven per cent. of these subjects ' failed ' the Arithmetic test in the first time limit, but this pattern was exactly reversed in the additional time limit, 77 per cent. now obtaining a 'pass' mark (to achieve this, of course, required a considerable improvement in score in the additional time). On the other hand, the low increase Matrices subjects not only tended to ' fail ' the Arithmetic test in the first time limit, but also showed a much smaller improvement in the additional time.

TABLE 3

RELATIONSHIP BETWEEN SCORE INCREASE IN PROGRESSIVE MATRICES AND PERFORMANCE ON ARITHMETIC TEST IN FIRST AND SECOND PERIODS.

Score Increase (Matrices)	" Passed " Arithmetic Test (scored 4+)		" Failed " Arithmetic Test (scored 4 —)	
	First 10′ Period	Second 10′ Period	First 15′ Period	Second 10′ Period
" Low " (Groups 0, 1, 2)	11	21	50	40
" High " (Groups 3, 4, 5, 6, 10)	6	13	20	13
" High " (Groups 7, 8, 13)	3	10	10	3

(4) *Personality Factors.*

Table 4 shows the results obtained with the Junior M.P.I., when the subjects are categorized into groups (0, 1, 2), (3, 4, 5, 6, 10) and (7, 8, 13) in terms of Matrices score-increase pattern. It is obvious that the groups do not differ significantly in respect of either neuroticism or extraversion. The neuroticism means are somewhat higher than those obtained by Furneaux and Gibson (1961) for a group of normal children ranging in age from 9–13 years ; while the extraversion means are very similar to theirs.

TABLE 4

RELATIONSHIP BETWEEN MATRICES SCORE INCREASE PATTERN AND JUNIOR M.P.I. SCORES.

Score Increase (Matrices)	Junior M.P.I. Scores			
	Neuroticism		Extraversion	
	Mean	S.D.	Mean	S.D.
Groups (0, 1, 2),	9·33	4·84	11·55	1·60
Groups (3, 4, 5, 6, 10)	8·00	4·88	11·80	1·00
Groups (7, 8, 13)	9·31	4·82	11·61	1·77

V.—DISCUSSION.

No elaborate discussion of these results is necessary. They confirm and extend previous results reported for university students with Progressive Matrices and objective tests and suggest very strongly that examinations with arbitrarily imposed time limits may lead to a serious underestimate of the potential level of a sizeable group of subjects. They do not, however, throw any light on the factors which may produce a tendency to work slowly but accurately in subjects whose level of ability is above average, except in-so-far as they suggest that neuroticism and introversion-extraversion may not play a predominant role. However, the implications for current systems of selecting students for further education may be important.

VI.—REFERENCES.

FURNEAUX, W. D., and GIBSON, H. B. (1961). A children's personality inventory designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol.*, 31, 204-208.

LYNN, R., and GORDON, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 194-203.

RAVEN, J. C. (no date). *Progressive Matrices* (1947), *Sets I and II*. Published privately by the Crichton Royal.

YATES, A. J. (1963). A further study of Progressive Matrices (1947). *Brit. J. Educ. Psychol.*, 33, 307-311.

YAYES, A. J. (1966). The relationship between level and speed on two intelligence tests. *Brit. J. Educ. Psychol.*

(*Manuscript received* 1*st July*, 1965)

RESEARCH NOTES

THE ASSESSMENT OF READING AND ARITHMETIC RETARDATION IN THE SCHOOL

By R. D. SAVAGE

(*Department of Psychological Medicine, University of Newcastle-upon-Tyne*)

AND D. J. O'CONNOR

(*Alnwick College of Education, Northumberland*)

A frequent problem in both the school and clinical setting is the assessment of reading or arithmetic retardation in a particular child. There is a general acceptance that a consistent relationship exists between I.Q. and reading and/or arithmetic ability. Little information, however, is available on the precise variation in arithmetic and reading ability which may still be considered normal in relation to a particular intellectual level. A step in the right direction came in the paper by Fransella and Gerver (1965), presenting valuable information from a clinical setting which related chronological age and intelligence to expected reading level. The problem of diagnosis in the school, however, requires economic screening methods applicable to groups of children.

The present research note gives further information on the relationship of reading and of arithmetic to intellectual level which derives from a four-year developmental study begun in the Northumberland schools (Savage 1966). Analysis of the first year data is of value in relation to the assessment and prediction of reading and arithmetic ability in 7-year-old school children. This age level is particularly important as early signs of reading or arithmetic retardation should and can be diagnosed in the school setting and remedied before the condition becomes extreme.

A sample of 419 Northumberland school children (average age 7·6 years, S.D. ·38 years) of both sexes from eight urban and rural schools were given (i) the Otis Quick Scoring Mental Ability tests (Form A or Form B 1936-54) randomly, (ii) the Watts Sentence Reading Test 1 (1965), and (iii) at random, either the Schonell Essential Arithmetic test or the Essential Problem Arithmetic (Forms A, 1965).

RESULTS AND DISCUSSIONS.

Table 1 presents the Means and S.Ds. and intercorrelations of the Intelligence, Arithmetic and Reading Quotients.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF THE TOTAL SAMPLE.

Measurement	Mean N=419	Standard Deviation	Intercorrelations 1	2	3
Performance Scale					
1. Intelligence Quotient	109·77	12·10		2·60**	·380**
2. Arithmetic Quotient	95·11	19·18			·400**
3. Reading Quotient	102·52	20·89			

** P< ·01.

The intercorrelations support the view that there are significant relationships between I.Q., A.Q. and R.Q. (r=P< ·01 in all cases).

TABLE 2

EQUATIONS.

Age	Regression Equations	Standard Error of Estimate
6·9—7·9	$R.Q.=30{\cdot}84+{\cdot}653\ I.Q.$ $A.Q.=49{\cdot}89+{\cdot}412\ I.Q.$	19·41 18·52

The above data will allow the prediction of *reading* and *arithmetic* levels *expected* for a given I.Q. (the standard error of measurement is consistent with that found by Fransella, *et al.* (1965) for a similar age group as far as the reading is concerned).

For example :

A 7-year-old boy with an I.Q. of 110 would be expected to have an R.Q. of

$$\begin{aligned}&30{\cdot}84+({\cdot}653\times110)\\&=30{\cdot}84+71{\cdot}83\\&=102{\cdot}67\end{aligned}$$

His expected *A.Q.* would be

$$\begin{aligned}&49{\cdot}89+({\cdot}412\times110)\\&=49{\cdot}89+45{\cdot}32\\&=95{\cdot}21\end{aligned}$$

Note.—The apparently low R.Q. and A.Q. scores in relation to the I.Q., are consistent with the differences in mean scores and intercorrelations in Table 1.

The prediction of the abnormality or significant retardation of any *obtained* reading or arithmetic scores can be measured by calculating the significance of the difference between the obtained and expected scores. For example :

The normality or otherwise of an obtained R.Q. 90 for the boy would be estimated as follows :

$$\begin{aligned}&\frac{\textit{obtained R.Q.—expected R.Q.}}{\text{S.E.}_{\text{est.}}}\\&=\frac{90-102{\cdot}67}{19{\cdot}41}\\&=\frac{-12{\cdot}67}{19{\cdot}41}\\&=-{\cdot}653\quad P>{\cdot}05\ \text{N.S.}\end{aligned}$$

One might reasonably expect that a boy aged between 6·9—7·9 years with an R.Q. of 90 and an I.Q. of 110 is not a significantly retarded reader.

Thus, one can also predict from the data presented, the normality or otherwise of an obtained reading or arithmetic quotient. It must be remembered, however, that the I.Q. test used to calculate these regression equations was a group non-verbal test.

This information should be of value, particularly in the normal school, in the screening or early diagnosis of backwardness in reading or arithmetic and may well be used in this respect.

REFERENCES.

FANSELLA, F. and, GERVER, D. (1965). Multiple regression equations for predicting reading age from chronological age and W.I.S.C. Verbal I.Q. *Brit. J. Educ. Psychol.*, 35, 86-89.

OTIS, A. S. (1963-64). *Quick Scoring Mental Ability Tests—Alpha Form A and B.* London : George G. Harrap and Co..

SAVAGE, R. D. (1966). Personality factors and academic attainment in junior school children. *Brit. J. Educ. Psychol.*, 36, 91-92.

SCHONELL, F., and SCHONELL, E. (1965). *Essentials of Diagnostic and Attainment Testing.* London : Oliver and Boyd.

WATTS, A. F. (1965). *Sentence Reading Test I.* London : Nat. Found. Educ. Res.

(*Manuscript received* 17*th March*, 1966)

THE SOCIAL AND EMOTIONAL ADJUSTMENT OF CHILDREN WITH IMPAIRED HEARING ATTENDING ORDINARY CLASSES

By BRIAN FISHER*

Since the second world war, there has been a considerable development of services for the detection and treatment of children with impaired hearing, who are pupils in ordinary classes.

Estimates of the number of school children affected by impaired hearing vary. but there is evidence that as many as 5 to 10 per cent may be so affected at any one time. The hearing loss involved in the majority of these cases is likely to be of conductive type, of moderate degree, but of fluctuating character.

It is well established that such moderate hearing losses, which are the most characteristic hearing problem in the ordinary school population, are not readily noticeable to teachers or to medical officers. Children with such hearing losses may appear to hear quite normally when spoken to directly yet they may be unable to maintain the necessary listening posture when they are back in their class seating positions ; they appear to be able to hear when they want to hear.

It seems clear that, in the everyday classroom, surgery or school clinic situation, there is a distinct tendency for hearing impairment in children, attending ordinary classes, to pass unnoticed by teachers and doctors or, where noticed, to be considered as of minor importance.

A situation very frequently arises where the audiologist in charge of the schools audiometric screening service informs the teacher or doctor that a certain child has a hearing impairment yet there may be little evidence of it to the teachér or doctor.

Are these apparently slight hearing impairments of any real consequence ?

There is an increasing conviction among audiologists that hearing impairments in children in ordinary classes, including the apparently slight impairments, constitute a threat to adjustment and to school progress. This investigation attempts to examine this conviction by studying a representative group of children with impaired hearing who are pupils in ordinary classes. Most of the children in the group have moderate hearing losses, but some of them have more marked losses.

The sample group is composed of forty-one boys and forty-two girls, ranging in age from 5·4 years to 16·0 years, with a mean of 10·1 years.

Twenty-seven of these children attend infant schools, twenty-four attend junior schools, twenty-eight attend secondary schools and four attend grammar schools.

The sample consists of children who had been kept under regular audiological review because of chronic hearing impairment. Their hearing losses, at the time of the investigation, range from 20·00 decibels (B.S.) to 63·75 decibels (B S.) with a mean of 37·80 decibels (B.S.). Fifty-one have conductive hearing losses, twenty-five have perceptive and seven have mixed hearing losses.

In three-quarters of the cases the onset of impaired hearing as reported by the parents, was during the child's first two years of life. Perhaps, surprisingly, 72 per cent. of the conductive cases fall into this category.

In very many cases, effective medical treatment for the impaired hearing is not secured until the child is discovered by the school's screening service. Many parents are found to have noticed their child's lack of quick response to sound, but to have attributed this to inattention rather than to impaired hearing. Many parents had, in the past, sought the advice of family doctors, but had been told that the child's hearing was clinically normal.

An examination of the adjustment in school of these eighty-three children by means of the Bristol Social Adjustment Guides shows them to have a high mean score on the Guides (mean score=11·9). This score denotes an abnormally high incidence of maladjusted behaviour in the group as compared with the available norms for the Guides.

---

*Based on unpublished thesis in degree of M.Ed. University of Manchester, Department of Audiology and Education of the Deaf.

As a check on the meaning of the results of the Bristol Guides, a control group was selected consisting of paired children, next on the class register and of the same sex as a child with impaired hearing. The mean Bristol Guide score for this control group is 8·1.

Fifty-three per cent. of the impaired hearing group and 72 per cent. of the controls are classified as normally adjusted on the Bristol Guides; 47 per cent. of the impaired hearing group and 28 per cent. of the controls are classified as having unsatisfactory adjustment.

The difference between the means for he two groups is statistically significant ($t=2{\cdot}497$, $df=163$, $P=<{\cdot}05$). The distributions of scores for the two groups are statistically significantly different ($\chi^2=6{\cdot}387$, $df=2$, $P=<{\cdot}05$).

An examination of the patterns of maladjusted behaviour is made. Among the thirty-nine children with impaired hearing, showing unsatisfactory adjustment, 33 per cent. show demonstrative patterns, 26 per cent. show withdrawing patterns and 41 per cent. show mixed patterns, where demonstrative and withdrawing symptoms occur in fairly equal measure.

Among the twenty-three control children showing unsatisfactory adjustment, 65 per cent. show demonstrative patterns, 22 per cent. show withdrawing patterns and 13 per cent. show mixed patterns.

These differences between the two groups of children as regards patterns of maladjusted behaviour are very significant statistically ($\chi^2=20{\cdot}323$, $df=3$, $P=<{\cdot}01$).

These results indicate that the withdrawing pattern is not the most characteristic maladjusted reaction pattern of this group of children with impaired hearing in ordinary classes; more of these children show demonstrative and mixed patterns.

In the control group, demonstrative maladjusted behaviour is most prevalent as would be expected from previous work.

The girls with impaired hearing are found to show more marked differences both in amount and pattern of maladjusted behaviour, compared to the control girls, than do the boys with impaired hearing compared to the control boys.

Several factors are examined to determine their relationship to the adjustment of the children with imparied hearing, as measured by the Bristol Guides.

A small but very significant curvilinear relationship is found between the Bristol Guide score and Cattell personality factor of Adjustment ($eta=0{\cdot}315$, $P=<{\cdot}01$). This very small correlation between two measures designed to identify maladjustment is considered to be due to the poor validity of the Cattell Questionnaires (E.S.P.Q., C.P.Q., H.S.P.Q.) for these children whose language development is so poor as to interfere with their understanding of the test questions.

This finding suggests the need for a careful appraisal of the validity of questionnaire and other test methods of a verbal kind in the assessment of the personality of children with impaired hearing attending ordinary classes.

The children with impaired hearing are found to have a normal mean and fairly normal spread of non-verbal ability grades on the Raven's Matrices. They are, however, substantially retarded on the Chrichton/Mill Hill Vocabulary Scales.

These results lead one to question the use of essentially verbal tests, like the Terman-Merrill, as the sole measure of the intelligence of children with a history of even slight hearing loss.

The children with impaired hearing are found to be substantially retarded in the basic school subjects.

Four factors most closely associated with the adjustment of these children, as measured by the Bristol Guides, are: non-verbal ability, chronological age, educational favourability of the home, and all-round school progress.

Adjustment in these children is thus closely connected with ability and attainment.

A curvilinear trend is apparent between adjustment and chronological age; whilst the older children, over 11 years of age, show the best adjustment, those in the age range 8–11 years, show poorer adjustment than those aged 5–8 years.

The findings suggest that impaired hearing may have a proportionately greater deleterious effect upon the adjustment and attainment of girls, who are pupils in ordinary classes, than boys. This requires further verification and examination in other studies.

The results, in general, indicate that the problems of impaired hearing, among children in ordinary classes, tend to be associated with a concomitant problem of emotional and social adjustment. This problem of adjustment is felt to constitute a serious threat to the child's ability to use his talents in a useful way in school. Without special assistance from audiologists there is a danger that impaired hearing, poor adjustment and scholastic failure may interact to produce retardation out of all proportion to the degree of hearing loss.

The need is made clear for the development of extensive audiological services to assist the child with impaired hearing in the ordinary classes and to promote the understanding of his difficulties and needs among parents, teachers and doctors.

(*Manuscript received 17th February,* 1966)

WHAT IS THE WEIGHT OF AN EXAMINATION PAPER?

BY JAMES M. THYNE

(*Department of Psychology, Jordanhill College of Education*)

SUMMARY. Examiners are often advised that, if one examination paper is to have twice the weight of another, it is only by the former's having twice the S.D. of the latter that the required effect will be achieved. But precisely what *is* the effect which is called 'having twice the weight'? And will doubling the S.D., in fact, achieve it? The answers call for a definition of Weight. Some established definitions are examined, and it is shown that, on these, doubling the S.D. will only rarely have the effect promised. A simple definition is proposed, which appears to be in accord with examiners' intentions, and which is such that a paper's weight is directly proportional to its S.D.

It is often affirmed, in texts on educational measurement, that when marks in several papers in an examination are combined, the relative weights of the papers' contributions to the total are directly proportional not, as is commonly believed, to their maximum possible marks, but to their standard deviations (e.g., Garrett 1939, Guilford 1936, Lindquist 1940, Vernon 1940). Expressions of this view have appeared recently in references to the Certificate of Secondary Education (e.g., Mather, France and Sare 1965, Secondary School Examinations Council 1964). There is, nevertheless, another view, also of long standing, namely, that while the papers' weights do depend to some extent upon their S.Ds., yet the weights are dependent upon other parameters also, particularly the intercorrelations of the contributing distributions (e.g., Gulliksen 1950, Sluckin 1956, Stalnaker 1938, Vaughn 1951, Wilks 1938). In the former view, a paper with an S.D. of 12 would have twice the weight of a paper with an S.D. of 6; whereas, in the latter view, this would not necessarily be so, and in most cases would not be so.

The difference in viewpoint may be of no practical significance when, as in the devising of a predictive test-battery, the marks in the total are to be compared with the marks in some external criterion. The coefficients (multipliers) to be applied to the various papers or sub-tests are mathematically determined so as to maximise the correlation between the criterion and the test-total; and if it were shown that Papers A and B had to be allocated 'weighting-coefficients' of 1 and 2, respectively, there would seem to be no practical objection to one's speaking of Paper B's having been given 'twice the weight' of Paper A. This would be understood to be but a conveniently short way of saying that this was the particular ratio (2 : 1) necessary for maximising the total/criterion correlation; and any argument as to whether the doubling of B's S.D. did, in fact, double some other thing, called 'its weight,' would probably be irrelevant. Nor would it be of more than academic interest to ask what this other thing, called 'weight,' might be.

These two questions do become of practical importance, however, in the case of most school, college and university examinations, where there is no external criterion with which the examination totals are to be maximally correlated. Here the weights are not computed, but are mathematically *in*dependent variables assigned at the outset in consequence of educational considerations, such as that Paper B is more relevant or required more work than Paper A. In this case, the weighting operations (such as doubling the S.D.) are performed upon the contributing distributions in order to achieve the previously specified weights *within the total itself.* Paper B would be given twice the S.D. of Paper A on the ground that this operation would weight the *total* twice as much with B as A. (Weighting to maximise a total/criterion correlation will also have effects upon the total; but in the non-criterion kind of examination it is within the total itself that the prescribed effect is to be achieved.) In this case, then, it is pertinent to ask whether giving Paper B twice the S.D. of Paper A will, in fact, give Paper B twice the weight of Paper A in the total.

That question is not properly answered by saying that doubling the S.D. *means* doubling the weight, for that merely makes 'weight' synonymous with 'standard deviation,' and so makes a tautology of the assertion that doubling the S.D. will double the weight; whereas, it is clear that most authors who make that assertion intend it to be a factual statement. Mather, France and Sare (p. 120) go so far as to say that "it can be shown mathematically." Whether it *can*, indeed, be shown is beside the present point, namely that, by those who make it, the assertion is intended to be treated not as an implicit definition of Weight (identifying it with S.D.) but as a theorem. It is from this point that the need of a definition of Weight becomes apparent.

That theorem, or any other, about the mathematical conditions governing a paper's weight is liable to proof; and this proof necessitates a mathematical definition of Weight, for, unless we specify Weight in mathematical terms, we cannot prove whether a paper's weight is directly proportional to its S.D. Putting it to the test is complicated, however, by so few texts' providing an explicit definition of Weight upon which proof could be based. Usually, they provide only illustrative examples from which a definition has to be inferred. Nor are all the implicit definitions the same, though two of these appear to be modal. (See Stalnaker.) One is that the weights of the contributing distributions are the (relative) degrees to which they contribute to the total *variance.* The other is that their weights are the (relative) extents to which they *correlate* with the total. These two views are now examined briefly, and it will be shown that in neither case is the theorem valid.

The variance of a total (σ_T^2), resulting from the combination of n distributions is:

$$\sigma_T^2=\sigma_1^2+\sigma_2^2+\ldots\ldots+\sigma_n^2+2(\sigma_1\sigma_2 r_{12}+\sigma_1\sigma_3 r_{13}+\ldots\ldots\sigma_{n-1}\sigma_n r_{n-1,n})\ldots\ldots(1)$$

which reveals that the total variance is a function not only of the variance of the contributing distributions, but also of these distributions' intercorrelations. To show more clearly the contribution of each distribution to the total variances the equation may be written (Vaughn) as follows:

$$\begin{aligned}\sigma_T^2=&\sigma_1(\sigma_1+\sigma_2 r_{12}+\ldots\ldots\ldots+\sigma_n r_{1n})\\&+\sigma_2(\sigma_1 r_{12}+\sigma_2+\ldots\ldots\ldots+\sigma_n r_{2n})\\&+\\&+\\&+\sigma_n(\sigma_n r_{1n}+\sigma_2 r_{2n}+\ldots\ldots\ldots+\sigma_n)\ldots\ldots\ldots\ldots\ldots\ldots(2)\end{aligned}$$

From these equations, it is clear that there is no case in which a distribution's contribution to the total variance is directly proportional to that distribution's S.D. There is, nevertheless, a special case in which a distribution's contribution to the total variance is directly proportional to (indeed *is*) its own *variance*, namely, when all the intercorrelations are zero. This special case scarcely merits consideration, however, because apart from its being out of accord with the theorem, it is unlikely that the intercorrelations of the papers in any one subject will be zero.

A related definition of Weight, suggested by (among others) Garrett, is that the weight of a distribution is its relative contribution, not to the variance, but to the S.D., of the total. He provides an illustration intended to support the view that (on this definition) a distribution's weight is directly proportional to its S.D. As Equation 1 reveals, however, this is true in only a special case, namely when all the intercorrelations are +1. (In that special case the total variance becomes the square of the sum of the S.Ds. and so the S.D. of the total is the sum of these S.Ds.) Garrett's illustration supports his thesis because in that illustration the intercorrelations *are* +1. Here again the special case scarcely merits consideration because examination papers generally do not yield perfect positive correlations.

When Weight is defined in terms of dispersion, then, the theorem that a paper's weight is directly proportional to its S.D. is untenable.

The second modal form of definition, that a paper's weight is directly proportional to its (relative) correlation with the total, is now examined.

If the S.Ds. of the n contributing distributions are $C_1\sigma, C_2\sigma, \ldots\ldots\ldots C_n\sigma$, then the correlation of any distribution ' p ' with the total is :

$$r_{pT} = \frac{\sigma(C_1 r_{p1} + C_2 r_{p2} + \ldots\ldots\ldots + C_p + C_q r_{pq} + \ldots\ldots\ldots C_n r_{pn})}{\sigma_T}$$

Similarly, the correlation of any other distribution ' q ' with the total is :

$$r_{qT} = \frac{\sigma(C_1 r_{q1} + C_2 r_{q2} + \ldots\ldots\ldots + C_p r_{qp} + C_q + \ldots\ldots\ldots C_n r_{qn})}{\sigma_T}$$

Accordingly, the ratio of the correlations of any two distributions with the total is :

$$\frac{r_{pT}}{r_{qT}} = \frac{C_1 r_{p1} + C_2 r_{p2} + \ldots\ldots\ldots + C_p + C_q r_{pq} + \ldots\ldots\ldots C_n r_{pn})}{C_1 r_{q1} + C_2 r_{q2} + \ldots\ldots\ldots + C_p r_{qp} + C_q + \ldots\ldots\ldots C_n r_{qn})} \quad \ldots .(3)$$

From Equation 3 it appears that this ratio does not equal $C_p : C_q$, except in the highly improbably case in which, in both numerator and denominator, the sum of the products is zero. (For this to happen, every correlation must be zero, or the sum of all the positive products must equal the sum of all the negative products. And the latter depends upon the values of the Cs as well as of the rs.)

In connection with this definition of Weight, two points made by Gulliksen are worth noting. While agreeing that there is a sense in which distributions correlating equally with the total may be said to have ' equal weight,' he points out that while this view enables us to define ' equal weight,' ' greater weight,' and ' less weight,' it does not permit us to speak of one distribution's having so many ' times the weight ' of another. The second point is that if some of the contributing distributions correlate highly one with another, these distributions will necessarily have similar correlations with the total, and so it will not be open to the examiner to assign a high weight to one of these distributions and a low weight to another of them.

On none of these definitions, then, is the theorem valid. We are left with two possibilities. We might accept one or other of the modal definitions, reject the theorem as invalid, and enunciate another in its place. This would, nevertheless, be somewhat unrealistic. As Equations 2 and 3 show, allocating specified weights necessitates the examiner's computing all the intercorrelations, a task few examiners would be prepared to undertake. Also, neither of these definitions appears to be in accord with what examiners generally *mean* when they speak about Weight. They wish to speak of one distribution's having ' so many *times* ' the weight of another whereas (as was noted) the ' correlation with total ' definition precludes that purpose. As for the ' dispersion ' definition, it is not clear how that concept relates to what examiners have in mind. The second possibility is that we might invent a different definition of weight which *would* be in accord with examiners' intentions and which *would* make the widely advertised theorem valid. Without such a definition, which would allow the examiner to see clearly the *effect* of doubling (etc.) a paper's S.D., and how that effect related to his purpose, the question, " Why are we weighting ? " is more than facetious paranomasia. A definition meeting these two requirements is now proposed.

The view upon which the proposed definition is based is an extension of illustrations used by Lindquist and by Vernon. Its crucial feature is that it focuses attention not upon the characteristics of whole distributions (e.g., variance and correlation), but upon what happens to specific candidates. Two of the candidates, call them U and V, are in a special tit-for-tat relationship: V's standing in Paper B is the same as U's standing in Paper A, while U's standing in Paper B is the same as V's standing in Paper A. The two candidates have equivalent pairs of standings; and the authors, taking the view that when Papers A and B have the same weight the two candidates will have equal totals, show that this will happen only when the two papers have equal S.Ds.

Now, in the absence of an explicit definition of Weight, the above demonstration is not a *proof* that equality of S.Ds. will give the two papers equal weights. What it does is indicate the *meaning* 'equal weight' has for the demonstration's authors. The differences in standing between U and V in Papers A and B are equal and opposite and, when the S.Ds. are equal, these two differences will cancel out. It is in this case only (they say) that the two papers will have 'equal weight.' We may deduce, then, that when the two papers do *not* have equal S.Ds., they would be said *not* to have equal weights, and so that one paper *would* be said to have '*greater* weight' than the other. This 'greater weight' is the result of one S.Ds. being greater than the other. Then precisely what *is* the effect of one paper's having (say) twice the S.D. of the other? To answer this question, let the z-scores (standings) of U and V in Papers A and B be u_A, u_B, and v_A, v_B, respectively. The difference between the two candidates' (raw) totals is:

$$D = \sigma_A (u_A - v_A) + \sigma_B (u_B - v_B) \ldots\ldots\ldots\ldots (4)$$

But the illustration makes $u_A = v_B$, and $u_B = v_A$, and so $(u_A - v_A) = -(u_B - v_B)$. Hence, when $\sigma_A = \sigma_B$, the difference in totals is zero, as in the original illustration. If, however, $\sigma_B = 2\sigma_A$, the difference will not be zero, for the difference in standings in Paper B, namely $u_B - v_B$, will be counted into the total difference twice as often as the difference in standings in Paper A, namely $u_A - v_A$, and the two quantities will not cancel out. The difference in total will now be in V's favour, instead of being zero; and since it is in Paper B that V beats U, it would seem sensible to say that it is Paper B which now has the 'greater weight.'

Further, although the original example referred to *equalities*, namely $u_A = v_B$ and $u_B = v_A$, yet all that has been deduced from Equation 4 could have been, deduced on the more general assumption that $(u_A - v_A) = -(u_B - v_B)$; that is, the assumption that the *difference* in standing between U and V in Paper A was the same as the *difference* in standing between V and U in Paper B. The essence of the illustration, then, is that, when one S.D. is m times the other, the z-score difference between the candidates in the one paper is being 'counted into' the total difference between them m times as often as is the z-score difference between them in the other paper. And it would seem appropriate to say that the total *difference* between the candidates is being 'weighted' m times as much with the former z-score difference as with the latter.

The example may now be further generalised to deal with *any* two candidates, U and V, whose z-scores in Papers A and B are u_A, u_B, and v_A, v_B, respectively. (Their z-score differences now are not necessarily equal and opposite.) Equation 4 is still applicable and shows that, when the S.Ds. are equal, the z-score difference between any two candidates in Paper A is counted into the total difference the same number of times as is the z-score difference between them in Paper B. But when $\sigma_B = m\sigma_A$, the z-score difference in Paper B is counted into the total difference m times as often as is the z-score difference in Paper A. And in this more general case also it would seem appropriate to say that the total difference between these two arbitrarily chosen candidates is 'weighted' with the A and B z-score differences in the ratio 1 : m. And this ratio is, of course, the ratio of the S.Ds.

The complete generalisation entails not two but n papers, having S.Ds. of $C_1\sigma$, $C_2\sigma, \ldots\ldots\ldots C_n\sigma$. Let any two candidates U and V have z-scores of u_1,

$u_2, \ldots\ldots u_n$, and $v_1, v_2, \ldots\ldots v_n$. The difference in total between U and V is :

$$D = \sigma \left[C_1(u_1 - v_1) + C_2(u_2 - v_2) + \ldots\ldots\ldots\ldots + C_n(u_n - v_n) \right] \ldots\ldots\ldots (5)$$

If all the Cs. are equal (equal S.Ds.), the z-score differences in the various papers are all being counted into the total difference the same number of times. If, however, $C_8 = mC_3$, the z-score difference between the candidates in Paper 8 is counted into the total difference m times as often as is the z-score difference in Paper 3. In general, the total difference between any two candidates is ' weighted ' with the contributing z-score differences in the proportions $C_1 : C_2 : \ldots\ldots\ldots : C_n$.

On this view, if the Papers 1, 2, ...p, ...n, have their z-score differences counted into the total $C_1, C_2, \ldots. C_p, \ldots. C_n$ times, then W_p, the weight of Paper p, may be expressed as the relative number of times its z-score difference, for any pair of candidates, is counted. Because of the way in which this view was arrived at, it will be obvious that the value of W_p is directly proportional to the S.D. of Paper p ; but to avoid tautology in the enunciation of the theorem the Cs. must be defined, as in this paragraph, not as the ratios of the S.Ds., but as the numbers of times the z-score differences are counted into the total difference.

Now, in order to *justify* the above definition, it was necessary to speak of the z-score *differences* being combined to form a total *difference*. But counting in these z-score differences C_p times is mathematically equivalent to counting in, C_p times, the z-scores themselves to form a total mark. We may, therefore, *state* the definition of Weight more simply by saying that the weight of a paper, in combination, is the relative number of times its z-scores are counted into the total marks. W_p, the weight of Paper p, is then given by

$$W_p = \frac{C_p}{C_1 + C_2 + \ldots\ldots\ldots + C_n} \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

where C_p is the number of times the z-scores in Paper p are counted into the total marks. On this definition, the weight of the total itself is unity, each of the contributing papers contributing a fraction. If there were three, papers, having Cs. of 1, 2, and 3, their respective weights would be 1/6, 1/3, and 1/2, which add up to unity and are in the ratios 1 : 2 : 3.

This definition appears to have two main advantages : it is reasonably simple and it legitimises the theorem. The latter point has already been made. It is simple in that its implications can be stated in a form easily understandable even by the examiner who is unversed in statistical niceties. Thus, to give Paper A twice the weight of Paper B is, in effect, to ensure that the candidates' standings in Paper A are counted into the total twice as often as are their standings in Paper B. The effect is identical to that of having Paper A sat twice with exactly the same results both times. (What are counted twice are, of course, the *standings* (z-scores) in Paper A, not the raw numerical marks). The ultimate criterion is, nevertheless, that the definition should be in accord with examiners' intentions. Being derived from their illustrations, it is presumably in accord with the intentions of Vernon and Lindquist ; but one cannot be certain that it is in accord with the intentions of school, college and university examiners generally, because it is rarely that their intentions are made explicit. When, however, it is demanded that one paper should be more ' important ' than another, the implication seems to be that in the ' more important ' paper a good mark should count for more, in the total, than should an equally good mark in the ' less important ' paper. And since that requirement can be met only by making a poor mark in the former count for *less*, in the total, than a poor mark in the latter, the implied requirement is equivalent to requiring that the *difference* between two standings in the former should count for more, in the total difference, than should the same difference between two standings in the latter.

Because the Cs. were so defined, assigning specific weighting-coefficients will certainly determine the relative number of times that the z-scores in the various papers will contribute to the total marks, and so will fulfil the examiners' intentions ; but the actual *values* of the total marks, and of the differences between them, do not depend upon these coefficients alone. The difference D, in Equation 5 is given in *raw* total scores; it is not the difference between any two candidates' z-scores in that total. To establish the *value* of D we should have to divide it by the S.D. of the total. The difference between two candidates' z-scores in the total is:

$$Dz = \frac{C_1(u_1 - v_1) + C_2(u_2 - v_2) + \ldots\ldots\ldots\ldots + C_n(u_n - v_n)}{\sqrt{C_1^2 + C_2^2 + \ldots\ldots C_n^2 + 2(C_1C_2r_{12} + C_1C_3r_{13} + \ldots + C_{n-1}C_nr_{n-1,n})}} \quad \ldots.(7)$$

Equation 7 shows that, when $(u_p - v_p)$ is positive, increasing the value of C_p will increase the value of the numerator (and increase the value of *raw* D), but will not necessarily increase D_z. And so unless the examiner takes the trouble to employ Equation 7, he cannot know precisely what effect weighting will have on a specified pair of candidates. The point is of no practical significance, however, because (we may assume) examiners will choose their weighting coefficients not in order to favour one preselected candidate at the expense of another, but with the intention that differences *generally* in Paper p should count for more in the total scores than should differences in other papers. It could be argued, indeed, that the examiner's ignorance of the effect upon individual examinees is desirable. He will know, in advance, the *kind* of candidate who will benefit from differential weighting, but he need not know who the favoured candidates are to be.

There is, nevertheless, one implication which might seem to detract from the definition's value. The (unknown) intercorrelations among the papers may be high, in which case weighting will make little overall difference to the total. To this possible objection there are two replies. First, the labour involved in weighting (as defined here) is very much less than the labour required to discover whether the intercorrelations are high enough to make weighting unnecessary; and even if it were unnecessary it would do no harm. True, weighting necessitates knowledge of the initial relationships among the S.Ds.; but these should be known anyway, because the papers will weight themselves even if the examiner fails to appreciate the point, and there is no excuse for combining marks when the S.D. relationships are unknown. The second reply is that although the intercorrelations may be high, yet if the number of candidates also is high, a few of these candidates may have significantly different standings in the different papers. The *overall* effect of weighting in such a case may be small, but may make a crucial difference for a small number of candidates, perhaps the difference between Pass and Fail. It would seem to be inequitable to sacrifice even a small number of candidates in the interest of statistical convenience.

REFERENCES.

GARRETT, H. E. (1939). *Statistics in Psychology and Education.* London: Longmans, Green and Co.

GUILFORD, J. P. (1936). *Psychometric Methods.* London: McGraw-Hill Book Co.

GULLIKSEN, H. (1950). *Theory of Mental Tests.* London: Chapman and Hall, Ltd.

LINDQUIST, E. F. (1940). *Statistical Analysis in Educational Research.* Boston: Houghton Mifflin Co.

MATHER, D. R., FRANCE, N., SARE, G. T. (1965). *The Certificate of Secondary Education.* London: Collins.

SECONDARY SCHOOL EXAMINATIONS COUNCIL (1964). *Examinations Bulletin No.* 3. *The Certificate of Secondary Education: An Introduction to Some Techniques of Examining.* Her Majesty's Stationery Office.

SLUCKIN, W. (1956). Combining criteria of occupational success. *Occupational Psychology.* Vol. 30, Nos. 1 and 2.

STALNAKER, J. M. (1938). Weighting questions in the essay-type examination. *J. Educ. Psychol.*, XXIX, No. 7.

VAUGHN, K. W. (1951). Planning the objective test. *Educational Measurement* (Ed. Lindquist, E. F.). Washington: American Council on Education.

VERNON, P. E. (1940). *The Measurement of Abilities.* London: University of London Press.

WILKS, S. S. (1938). Weighting systems for linear functions of correlated variables when there is no independent variable. *Psychometrika*, Vol. 3, No. 1.

(*Manuscript received* 16*th June*, 1966)

AN INVESTIGATION OF THE CRAIG READER COURSE

BY K. D. DUNCAN
(*Army Personnel Research Establishment, West Byfleet*)

AND MAJOR R. FOSTER
(*Army School of Education, Beaconsfield*)

INTRODUCTION.

The distinctive feature of the Craig Reader Course, as of the Harvard Course[9, 10] on which many reading improvement courses in this country have been based[11], is a device—the Craig Reader—which paces reading by briefly exposing successive groups of words in a narrative. Unlike the Harvard films, the Craig Reader is intended for use as an individual training device.

Recently, McLaughlin has reported that rechnical college students, who completed the Craig Reader Course, failed to show a significant increase over a control group in the mean amount they were able to recall of Building Research Station Digests, which they had previously read for 5 minutes · this, despite the fact that the average amount they had read within the time limit was significantly increased.[8] This finding is disappointing in view of (i) the achievements of other reading improvment courses (Poulton's survey found increased mean comprehension scores in thirty-three out of forty groups and increased mean reading speed in all of them[11]) and (ii) the advantage, on the face of it, of the individual trainee being able to change the Craig Reader pace, from lesson to lesson, by an amount adjusted to his own performance.

THE PRESENT INVESTIGATION.

(i) *Subjects.*

Subjects were soldiers aged between 18 and 20 years, attending a fourteen-week course to bring them up to the educational standard required for entry to the Royal Military Academy, Sandhurst. All had either passed or were candidates for the Regular Commissions Board. An experimental group consisting of the first thirty-six soldiers to arrive was assigned to the Craig Reader Course. A further intake of twenty solidiers, who arrived fourteen days later served as a control group.

(ii) *Experimental Procedure.*

The experimental group completed the twenty lessons of the Craig Corporation's " Advanced Reading Course A "[4] during ten days of September, 1965, i.e., before the arrival of the control group. (The exercises and the device have been described by McLaughlin.[8]) There were two 1½ hour sessions daily, one in the morning and one in the afternoon, except Wednesday and Saturday afternoons, and Sundays, i.e., fourteen sessions in all. Subjects worked at the films and printed passages independently and at their own pace. As only one Craig Reader with associated materials was available for every three subjects, the actual time spent on the various course exercises was considerably less than the duration of the experimental sessions. The average time so spent is estimated at 14 hours R.A.E.C. officers supervised the sessions and regulated the subjects' progress from lesson to lesson so that all completed the final lesson on the same day.

(iii) *Testing.*

Two tests are supplied with the course, each consisting of a passage of approximately 2,000 words and ten multiple choice questions on its content. The tests, which may be referred to as A and B, are said to be equivalent for use either as an initial or a final measure of reading performance. Nevertheless, as a precaution, half of each group took A as an initial test and B as a final test; the remaining ubjects took the tests in the order B, A. The experimental group was tested initially on the Saturday before the Monday on which the experiment began, and finally, on the day immediately after the last lesson of the course. The control group took both tests on the same day.

Black's Comprehension Test for Training College Students[5] was administered to all subjects two weeks after the experimental group completed the course. The time allowed was reduced from the usual 45 to 30 minutes in view of the expected improvement in reading performance of the experimental group.

On the occasion of the administration of Black's Test, all subjects read a pamphlet of 4,200 words on German Reunification for 15 minutes. This was immediately followed by an untimed test consisting of twenty-five open-ended questions requiring short answers on the content of the pamphlet. A marking guide for the test was prepared. As scores on this current affairs reading test (CA1) were rather low, and in view of the desirability of obtaining test-retest data, it was readministered six days later, following a subsequent reading of the German Reunification pamphlet, lasting 30 minutes (CA2).

In addition to the foregoing tests, all subjects completed test AH4.[6]

RESULTS.

TABLE 1

CONTROL GROUP MEANS AND ASSOCIATED STANDARD ERRORS IN BRACKETS FOR THE TWO CRAIG TESTS.

Test	Reading Speed (words per minute)	Comprehension (questions correct/10)
A	354·4 (24·85)	8·5 (0·28)
B	347·7 (22·64)	8·4 (0·25)

(i) *Craig Corporation Tests.*

In the case of the control group, the difference in difficulty between Tests A and B appears to be negligible in terms of either reading speed or comprehension (see Table 1). The product moment correlation between w.p.m. on tests A and B was 0·79 and between comprehension scores 0·30; these correlations and the data in Table 1 are based on twenty subjects. In computing the means and standard errors shown in Table 2 below, scores on the Craig Tests were pooled, regardless of which test (A or B) they were based on.

(ii) *Current Affairs Reading Test.*

The test-retest coefficient (six days delay) was 0·71 for this test. The coefficient is based on the scores of both experimental and control groups, a total of forty-eight subjects. The original fifty-six subjects were reduced to forty-eight between the end of the reading course and the completion of testing, two experimental group subjects and six control group subjects being found unsuitable by the Regular Commissions Board and returned to their units during the experiment.

(iii) *Comparability of the Two Groups.*

The main results of the experiment are summarised in Table 2. Data for the eight subjects lost during the experiment have been discarded. The AH4 scores of the two groups are comparable, as are the initial Craig Test comprehension scores, but the initial reading speeds obtained by the Craig Test differ significantly ($p<0{\cdot}01$). However, this variable bears little relation to post course performance on Black's Test ($r=0{\cdot}16$), CA1 ($r=0{\cdot}15$), CA2 ($r=0{\cdot}09$) or the final Craig Test

TABLE 2

MEANS FOR COMPARISON AND POST-COURSE TESTS WITH ASSOCIATED STANDARD ERRORS IN BRACKETS.

Test	Experimental Group (n=34)	Control Group (n=14)
Comparison Tests :		
AH4 : Part 1	44·7 (1·20)	44·6 (0·93)
Part 2	49·2 (1·77)	44·9 (2·51)
Initial Craig Test :		
w.p.m.	233·6 (9·72)	299·3 (13·64)
comprehension	9·1 (0·22)	8·6 (0·31)
Post-Course Tests :		
Final Craig Test :		
w.p.m.	488·4 (21·03)	336·6 (15·76)
comprehension	8·5 (0·26)	8·3 (0·29)
Black's Test	28·2 (1·79)	21·1 (3·00)
Current Affairs Reading Test :		
C.A.1	8 9 (0·57)	7·4 (1·15)
C.A.2	14·0 (0·62)	12·0 (1·24)

comprehension score $(r=0{\cdot}02)$. Initial reading speed is significantly correlated with final reading speed obtained by the Craig Tests $(r=0{\cdot}65)$, so, had the two groups been more closely matched on this variable at the outset, the final difference between them would probably have been even greater than that observed.

(iv) *Differences in Post-Course Performance.*

In terms of reading speed as measured by the Craig Test, the experimental group is statistically significantly superior, at the end of the course, to the control group $(p<0{\cdot}01)$, but the magnitude of this difference (45 per cent.) may be an underestimate in view of the initial difference between the groups on this variable. The experimental group improved significantly $(p<0{\cdot}01)$, approximately doubling reading speed between the initial and final tests, whereas the improvement of the control group is not statistically significant. The initial and final speeds of the experimental group agree well with those reported by McLaughlin.[8] In neither group was the fall in comprehension scores statistically significant. Scores on Black's Test indicate a statistically significant difference of 33 per cent. in favour of the experimental group $(p<0{\cdot}05)$. As regards CA1 and CA2, the data were subjected to analysis of variance, and the overall difference between the two groups (19 per cent.) was statistically insignificant $(p=0{\cdot}11)$.

DISCUSSION.

The rationale of the Craig Course appears to be to increase reading speed while maintaining a ' constant ' level of comprehension as measured by the tests on the film and booklet passages. Clearly, the less sensitive the tests are, the more dramatic will be the increases in reading speed which are possible while still maintaining 'good' test scores. Thus, to take increased reading speed, as measured by the Craig Tests, as an index of the effectiveness of the course could be hazardous in view of the implicit assumption that comprehension has been maintained at a constant level. It might be concluded, for example, that the experimental group in this and McLaughlin's study[8] had doubled their reading efficiency.

Such an evaluation of the Craig Course appears to be advocated by the literature supplied with it. On the other hand, the scores on Black's Test, in which time is held constant and amount comprehended is measured, yield a smaller, but perhaps more defensible, estimate of the improvement in reading attributable to the course (33 per cent.).

Most ventures into improving reading are restricted to the package courses available. This is regrettable since the exercise material may or may not be suited to the needs of the subjects and the importance of mechanical devices, which are frequently incorporated in the package, is questionable[5, 7]. Furthermore, devices which present text through a 'moving window' in order to pace reading, may introduce other difficulties for the subject if the window is small; laboratory evidence indicates a substantial increase in errors as window size is reduced.[12] (The line of print exposed by the Craig Reader increases from the beginning to the end of the course; the range is from three to eight words or twenty to forty-five letter spaces.)

Anderson, in a small exploratory study, found little difference in reading speed over the first four lessons of the Craig Course when presentation of the filmed passages by the Craig Reader was replaced by self-paced reading of the same passages printed on sheets of paper.[1] On the other hand, McLaughlin, whose subjects were volunteers, despite having to attend sessions in the early morning, the lunch hour and at night, concluded that this high motivation was to some extent attributable to the 'fascination of using the machine.'[8]

Pressing the subject continually, to increase his reading speed may not be the only or the best way to exploit the Craig Reader. It seems that good comprehenders have apparently learned to adjust their rate of reading to the difficulty of the passage, whereas poor comprehenders tend to read at a constant rate.[3] If adequate knowledge of results in the form of sensitive comprehension tests could be provided, the Craig Reader might be used to train the subject to recognise his optimum reading rate for different sorts of material by adjusting the speed control up and down.

Efficient reading has always been important in many jobs and in most education and training. Its importance in the latter field is probably even greater with the introduction of programmed instruction which in practice, if not in principle, is very dependent on the presentation of printed material. Thus, any promising method of improving reading deserves study. In the case of the Craig Reader, further investigations are desirable both with and without the machine, with different sorts of reading matter as exercises, and with supporting comprehension tests sensitive enough to inform the subject when he is increasing reading speed at the expense of gross losses in comprehension.

REFERENCES.

[1] ANDERSON, J. D. (1965). Ministry of Defence, Science 4 (R.A.F.). *Personal Communication.*

[2] BLACK, E. L. (1962). *Manual of Instructions for Comprehension Test for Training College Students.* London: National Foundation for Educational Research.

[3] BLOMMERS, P., and LINDQUIST, E. F. (1944). Rate of comprehension of reading: its measurement and its relation to comprehension. *J. Educ. Psychol.*, 35, 449-473..

[4] CRAIG RESEARCH, INC. (1961). *Advanced Reading Program A.* Los Angeles: Craig Corporation. Ashford, Middlexes: International Tutor Machines, Ltd.

[5] GLOCK, M .D. (1949). The effect upon eye movements and reading rate at the college level of three methods of training. *J. Educ. Psychol.*, 40, 93-106.

[6] HEIM, A. W. (1955). *Manual for the Group Test of General Intelligence AH4.* London: National Foundation for Educational Research.

[7] JONES, D. H. (1965). Trianing industrial executives in reading: a methodological study. *J. Appl. Psychol.*, 49, 202-204.

[8] MCLAUGHLIN, H. (1964). Research Report. *New Education* 1, 1, 35-37.

[9] PERRY, W. G., Junr., and WHITLOCK, C. P. (1948). *Harvard University Reading Films, Series Two.* Cambridge, Mass.: Harvard University Press.

[10] PERRY, W. G., Junr., and WHITLOCK, C. P. (1948). *Selections for Improving Speed of Comprehension.* Cambridge, Mass.: Harvard University Press.

[11] POULTON, E. C. (1961). British courses for adults on effective reading. *Brit. J. Educ. Psychol.*, 31, 128-137.

[12] POULTON, E. C. (1962). Perpheral vision, refractoriness and eye movements in fast oral reading. *Brit. J. Psychol.*, 53, 409-419.

(*Manuscript received 17th March, 1966*)

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

A SOCIOMETRIC COMPARISON OF WELL-BEHAVED AND BADLY-BEHAVED CLASSES

By J. I. ROSE

(*Summary of Theses presented in part fulfilment of the requirements for the degree of M.Ed. in the University of Aberdeen*)

AIM.

To compare the morale, preferences, dislikes and sociometric structures of well-behaved and badly-behvaed classes in secondary schools, and to examine the behaviour traits and attitudes of the pupils who held high status in each type of class.

SUBJECTS

Six pairs of secondary school classes, matched for size, age and sex (four pairs boys', two girls' classes), each pair consisting of one well-behaved and one badly-behaved class, as judged by teachers' ratings. The total number of pupils involved was 314.

METHOD.

Pupils completed Fitt's Attitude to School Questionnaire, an eleven-item sociometric questionnaire, a modified Tryon ' Guess Who ' questionnaire, and ratings on class happiness and behaviour and their own behaviour. Juvenile Court appearances were also recorded. Teachers rated classes for morale (specifically defined) and listed the most co-operative, most troublesome, most preferred, least preferred, most popular and least popular pupils. Attendances over a twelve-week period were also recorded.

RESULTS.

The validity of the teachers' ratings of the classes as well-behaved and badly-behaved was checked against the objective data. Three hypotheses were tested.

(i) Badly-behaved classes would display characteristics of lower morale than well-behaved classes, in that they would be less integrated, would be less able to work together, would show less friendliness to each other, would be less likely to function as a unit and more likely to act as a collection of individuals. On this hypothesis, the problem of the badly-behaved class should be seen more in terms of morale than of discipline, and more attention should be given to raising the morale of the class. Evidence was conflicting, though this may be because the sociometric techniques are still crude and poorly developed.

(ii) Teachers of the badly-behaved classes would be less aware of the sociometric structure of such classes than teachers of well-behaved classes, and the teachers' preferences and dislikes among the pupils would be less in agreement with those of the pupils in badly-behaved classes than in well-behaved classes. This was confirmed, but in both types of classes there was a fair degree of agreement between teachers' and pupils' preferences.

(iii) The pupils in badly-behaved classes would select different behaviour traits for the pupils they selected as having high status, in contrast with the traits favoured by pupils in well-behaved classes. This was not confirmed. High-status pupils of both types of class were shown to possess traits characterized by values in keeping with the school (able to work quietly, tidy, active in games, good-looking, happy, ambitious), while the opposite was true of low-status pupils (attention-seeking, unkempt, fight-seeking, unhappy, not persevering).

It is concluded that the problem of badly-behaved classes might best be met by seeking to foster the unity of the class and to create situations which encourage the high-status pupils to make use of their influence.

THE EFFECT OF TRANSFER ON THE LEVEL OF CHILDREN'S ADJUSTMENT TO SCHOOL

By W. F. MURDOCH

(*Summary of Thesis presented in part fulfilment of the requirements for the degree of M.Ed. in the University of Aberdeen*)

AIM.

To examine the extent to which children experience difficulties in transfer from primary to secondary education, and the background factors associated with the occurrence of such difficulties.

SUBJECTS.

Two-hundred-and-eighty-nine boys and 263 girls from six primary schools from west-end, council estate and east-end housing areas in Aberdeen, who entered secondary school in January and August, 1965.

METHOD.

In December, 1965, after the children had spent three terms (January intake) or one term (August intake) in secondary schools, the children wrote a half-hour essay on instructions to " tell the story of your first few weeks in this secondary school, comparing it with your primary school. Think back and write about your feelings when you first started at this school and about your feelings now you have been here longer." Further suggestions were given and four specific questions on adjustment were asked at the end. The essays and the questions provided the basis of a classification (the reliability of which was checked by a re-classification of fifty scripts by a second marker) of the pupils to distinguish those who had encountered problems of adjustment and to assess the success of adjustment. Data on the children's verbal reasoning test scores, social class, sex, age, height and weight, were available from a parallel investigation of this group. The relationship of adjustment to these background data was examined.

RESULTS.

(i) 60 per cent. of the pupils reported some problem of adjustment, and 7 per cent., who declared that they preferred their primary school to the secondary school, were judged to have failed to regain the level of adjustment to school which they had attained before transfer.

(ii) As there was no significant difference between the proportions of those successfully adjusted in the January and August intakes, it would appear that for the majority adjustment is a rapid process.

(iii) No relationship was found between adjustment and the children's age, height or weight. There was a tendency for more intelligent girls to report less success in adjustment.

(iv) Children from social classes I and II showed more success in adjustment in senior secondary schools than those from (iv) and (v). No difference between these groups was found in adjustment to junior secondary schools.

(v) Difficulties of adjustment were related to : disapproval of changing classrooms ; disapproval of school rules and punishments ; failure to make new friends ; unpleasant treatment by older children ; dislike of having several teachers instead of one ; failure to find the work more interesting than in primary school.

AN INVESTIGATION INTO THE DIFFERENCES BETWEEN CHILDREN IN THE FORMATION OF CONCEPTS OF TOWN AND COUNTRY

BY JENNIFER F. MARSH

(*Summary of Thesis presented in part fulfilment of the requirements for the degree of M.Ed. in the University of Aberdeen*)

AIM.

To see whether a rural environment helps a child to develop the concepts of town and country, as a test of Piaget's explanation of the formation of these concepts.

SUBJECTS.

Two groups, each of forty-two 9-year-old children, matched for score on N.F.E.R. Sentence Reading Test 1, and socio-economic status on the Registrar-General's scale, one group from an Aberdeen city school, and one group from surrounding rural schools.

METHOD.

(*a*) By individual questioning, " What is Aberdeen ? Where is Aberdeen ? What is Scotland ? Where is Scotland ? What is Great Britain ? " Answers scored on basis of four stages of grasping the Aberdeen-Scotland-Britain relationship.

(*b*) By a spatial test involving cards and buttons, to be arranged to represent the Aberdeen-Scotland-Britain relationship (after a demonstration trial to show the pupil-room-school relationship).

RESULTS.

Rural children at this age are significantly more advanced than urban children in developing the concept of a town as a unit within a country, and in the development of spatial concepts of the part-whole relationships of Aberdeen-Scotland-Britain. A disturbingly high proportion of children showed confusion over the spatial concepts involved. The results are interpreted in terms of the daily experience of travel of rural and urban children.

CRITICAL NOTICE

LIAM HUDSON (1966). *Contrary Imaginations.* London: Methuen and Co., pp. 151+30, price in the U.K., 25s. nett.

BY F. V. SMITH
(*University of Durham*)

In one important sense, the so-called intelligence and aptitude tests are a substitute for time. If it were possible for all aspirants to spend sufficient time in a university or comparable institution to discover some of their abilities, the activities they liked best and whether or not they could cope with the work, in short, to regard the university as an adventure, many of the problems of selection would be greatly reduced; but administrators are often in a hurry and in the past have sought advance knowledge of the potentialities of the young, thereby launching a complex psychological problem and creating a situation fraught with implications for social and educational philosophy. Appropriately enough, some of the earliest programmes of mental resting, by Binet and Simon and the use of the alpha and beta tests by the American Army, were conceived in emergency situations and did provide useful, if coarse, answers in the special circumstances.

Usually a person, be he child or adult, who records an I.Q. of 120 on one of the so-called intelligence tests can be expected to reveal a general level of behaviour and competence very different from another person who records an I.Q. of 70; but correlations between intelligence test scores and academic success, particularly at higher levels of education, even when corrected for uni-variate selection, are frequently not impressive and naturally, there are some striking individual cases of discrepancy. This is an old and familiar story, re-examined and illustrated in the extensive studies of Sanders (1948), Garrett (1949), Stuit (1949), Hohne (1951) and indirectly re-stated in the more recent studies of Roe and of MacKinnon who showed that above a certain minimum level of test score, which in some cases can be surprisingly low, success and even eminence in difficult and creative activities esteemed in Western culture is not indicated by test performance—or, as Dr. Hudson, with typical succinctness, states, "A mature scientist with an adult I.Q. of 130 is as likely to win a Nobel Prize as one whose I.Q. is 180," or again, "A boy's I.Q. is of little use, if you are faced with a formful of clever boys."

The investigation of many discrepancies among school populations and analysis of individual cases has led Dr. Hudson to be disappointed profoundly with conventional intelligence tests. With something of the fervour and the charm of the evangelist newly emerged from the wilderness, he writes very crisply in admonition of their long continued use and the attempts at theoretical under-pinning. There is scant mention of studies involving multiple regression, which assume that intelligence test scores are only one variable in the problem and of some of the side benefits, such as improved scaling methods and the application in other sciences of methods of factor analysis developed by psychometricians.

Close examination of the performance of individual boys further led him to take up qualitative considerations along lines suggested by Getzel and Jackson's distinction between the 'High I.Q.' and the 'High Creative.' For the person who records his best performance on the rigorous and exacting deductions, analogies and calculations which abound in the conventional intelligence test, he uses the term 'converger' because normally, with this type of item, there is only one answer and the candidates, reasoning is required to converge thereon. Rather different are the 'open-ended' tests which call for inventive suggestions and thus admit of many answers. Dr. Hudson prudently refrains from calling them 'creativity tests' to avoid implications of the kind deriving from the phrase 'intelligence test' and because of inadequate evidence that they are valid tests of creativity.

However, the person who records his best performance on these tests is termed a ' diverger.' With a school-boy population, a suggested distribution is : ' extreme divergers ' 10 per cent. ; ' moderate divergers '—20 per cent. ; ' all-rounders '—40 per cent., and moderate and extreme ' convergers ' —20 and 10 per cent., respectively. The classification refers only to bias, so that it would be possible for a ' converger ' to have a higher score on the open-ended tests than a ' diverger.'

Examples of open-ended tests are indicating ' Uses of Objects ' and ' Meanings of Words,' ' Drawing to illustrate a Title,' comments on ' Controversial Statements,' ' Written Autobiography ' and the candidate's ' Rating of Personal Qualities.' In indicating ' Uses of Objects,' candidates are invited to suggest uses for such objects as barrel, paper clips, tin of boot polish, brick and blanket. As might be expected, the number and novelty of the suggestions varies from person to person, with the ' divergers ' supplying more of the occasional, luscious suggestions. The central finding was that most of the boys who specialise in Arts subjects (history, English literature and modern languages) and who did poorly on I.Q. tests, were much better at open-ended tests and could be classified as ' divergers.' Specialists in physics, mathematics, chemistry and classics, include more ' convergers,' while biology, geography, economics and general arts courses are claimed to attract ' convergers ' and ' divergers ' in about the same proportion. Other observations are that ' divergers ' are more likely to suggest rare uses for objects, to suggest, and elaborate upon amusing aspects, to include references to people, and to suggest uses involving discomforture of others and even violence ; but when the ' convergers ' do refer to violence, it is more likely to border on the morbid and ghoulish as if their more tightly controlled defences had suddenly disintegrated.

Following a cue indicated in a study by McClelland of American scientists, which suggested an association between eminence and a radical protestant or puritanical home background, Dr. Hudson speculates that such a home in which parents are more insistent on correct behaviour and the mastery of impersonal, practical skills is more conducive to 'convergence' and conventional scientific rigour than a more permissive and emotionally warmer home atmosphere which would tend to favour the development of children as ' divergers.' For these, Dr. Hudson appears to have some sympathy because they are deemed more likely to do unexpected things and may be stunted by conventional systems of schooling.

Now this approach offers highly seasoned fare indeed. It has many of the fascinations of depth psychology and raises not only vexed issues involving social and educational values, but also the spectre of waste of the precious national assets of scientific and artistic creativity. It is a rich field for polemical writing for which, one suspects, the author has a certain relish and flair. He modestly declares his intention (p. 18) of " describing research which has no immediate practical value " and confesses to " a fear that the book will seem socially useful when it is not." His aim is " to use the system as it now stands as a source of insight into the ways in which children feel and think."

Any sincere attempt to enhance the understanding of qualitative aspects of mental life commands the immediate respect of all psychologists ; but there will be many of them, despite the chiding they receive in the book, who will ponder whether or not, and on rather tentative evidence, Dr. Hudson has started yet another typology, which may run its course through the literature, and eventually merit much of the admonition he has applied to the mental testing movement of the last half century. F.V.S.

BOOK REVIEWS

ADLER, SOL. (1966). *A Clinician's Guide to Stuttering.* U.S.A.: Charles Thomas, pp. xiii+175, no price.

On the cover of this book we are told that it is written for " the student speech and hearing major," and that it is " . . . a comprehensive introduction to the whole problem of stuttering." If the latter were the case, then this book might serve a useful purpose in this country, but unfortunately, it is not a very good introduction and it is certainly not comprehensive. For instance, there is no mention at all of the many problems involved in the measurement of stuttering, nor of the considerable literature on the relationship between delayed auditory feedback and stuttering.

One of the things that makes this book rather annoying to read, is the many categorical statements based on very little evidence. For instance: " stuttering often fashions and moulds a neurotic personality," and " the stutterer has become conditioned to persons as well as to situations." There is also a constant alternation between the language of psychoanalytic theory and learning theory which does not make for easy reading.

The main emphasis of this book is on the author's own views on the treatment of stuttering, particularly of the stuttering child. There is a good section on how the teacher might help such a child and on the advice that could be given to parents. Of particular interest is a section on stuttering in the mentally retarded, a topic on which very little has been written. This book may, therefore, be of some value to the teacher, but is unlikely to provide the sort of information required by the psychologist interested in treating stuttering or carrying out research in this area.

FAY FRANSELLA.

ATKINSON, R. C., BOWER, G. H., and CROTHERS, J. (1965). *An Introduction to Mathematical Learning Theory.* New York: John Wiley and Sons, pp. xiii+429, 75s.

With recent trends to develop and apply mathematical methods in the social sciences, interest has shifted from mere statistical treatment of data to the increasing use of methematical models in theory construction and evaluation. However, with this emphasis some of the social scientists fear that they would lose their science in the complex elaborate maze of mathematical derivations. This fear, I think, is allayed to some extent by the authors of this book in that it beautifully focuses its mathematical attention on the psychology of learning, an important area of experimental psychology which with its own unique methods offers a useful source of systematically collected observations to guide an empirically oriented theory. Beginning students and research workers often shun the books on mathematical psychology, though it has now become almost an essential for researchers to get thoroughly acquainted with quantitative methods. As one goes through the pages of this work, it becomes apparent that the present authors have fully kept in view this limitation and tried to make this book an introductory primer for those who possess only a modest mathematical training. It is a psychology book, not a mathematics book, though it illustrates basic mathematical methods by specific case studies of selected learning theories in psychology.

The book presupposes on the part of the reader a knowledge of college algebra (occasionally including elementary understanding of matrix algebra and calculus) and an elementary course in statistics (covering some probability theory). The authors aim to prepare the readers with this background to assimilate the more advanced developments in mathematical psychology with the help of the models and topics discussed in the book. The main research topics treated are: concept identification, paired-associate learning, choice behaviour, signal detection and probability learning, avoidance conditioning, stimulus sampling theory, social and economic interactions, and parameter estimation. The first is an introductory chapter to stress the role of mathematics in theory construction. Second and third chapters review notions of elementary probability theory and also introduce Markov chains and linear difference equations. The remaining chapters can be pursued somewhat independently, according to the reader's interests.

No doubt, the book covers a relatively wide range of topics and models ; yet the authors do not claim to survey all of the learning models nor to survey exhaustively all of the empirical evidence related to models discussed. In brief, the book gives a fairly complete picture of how a mathematical model for learning works in action covering its formulation, development and experimental application. It needs be admitted that the book enables any reader acquainted with the psychology of learning and statistics to master the fundamentals of mathematical learning theory. Authors' own research interests and data give the book a valuable touch of empirical evidence and will evidently encourage the readers to go beyond the topics to construct models of their own areas of specialization after they have mastered some of the techniques illustrated in the book. A. S. PATEL.

BARBE, WALTER B. (Ed.) (1965). *Teaching Reading: Selected Materials.* New York : Oxford Univ. Press, pp. xiii+430, paperback, no price given.

Fifty-three articles from various American publications have been reprinted here. They are grouped into sub-sections, each of which is introduced by the editor. The whole of the article is presented in every case and the dates of the original publications range from 1926 to 1964, the majority having first appeared in the 50's or 60's. With the exception of Sir James Pitman and Professor P. E. Vernon, all the writers are American. Included are Gray, Witty, Gates, Anderson, and Dearborn, Dale, Monroe, and Bear.

The main theme is concern over the improvement of 'Johnny's' reading. Article after article deals with aims in reading programmes, and possible approaches. Oral and silent reading, phonics and look and say, modern systematic and opportunistic methods, individualised reading and homogeneous group teaching are among the topics which are fairly extensively covered. Developmental programmes are discussed and newer issues such as i.t.a , the contribution of linguistic and pre-school learning are not forgotten. The section on word perception and meaning, including articles on vocabulary control and the acquiring of meaning from context, is well written and interesting.

It is to Barbe's credit that a balance is achieved in all this. Some undue claims are made, of course, but on the whole the impression is of a cautious and sensible widening of knowledge. It is also to his credit that he himself has written up an experimental approach which did not work and has included articles which come out against his own convictions about method.

It is possible to be highly critical of the American predilection for assembling collections of this kind. What can be said in favour of this particular one ? Its main contribution is that it makes immediately available articles which it would be laborious to pursue in the original publications, expecially for British readers. The majority of the articles carry useful references. The introductory statements to each section summarise the main issues neatly. On the other hand, some areas are rather inadequately dealt with. Remedial education suffers here. Re-printing the whole of the article each time does produce repetition. The sections need some discussion at the end as well as the beginning.

Barbe himself sees this book as having value in teacher training programmes. American students would surely have benefited from the inclusion of some of the valuable findings by British workers in the field. The almost complete exclusion of the latter detracts from the book's usefulness in initial training courses in this country. It would, however, provide interesting discussion material in advanced courses for practising teachers and in in-service training. ALICE F. LAING.

Clarke, R. B., Coladarci, A. P., and Caffery, J. (1965). *Statistical Reasoning and Procedures*. Columbus, Ohio: Charles E. Merrill Books, pp. ix+390, 56s.

The purpose of this introductory text is to help non-mathematicians in psychology and education to "develop an intuitive understanding of statistical reasoning." Believing that this can only be achieved with reference to applications of specific statistical procedures, the authors limit themselves to simple exposition of a few basic procedures.

This bold omission of several important areas usually covered in introductory texts, notably analysis of variance, non-parametric statistics other than chi-square, conditional probabilities, statistical decision theory, multivariate analysis and the like, results in a clear and well organised exposition of the formulae and computational steps of those few procedures covered. Thus, the first eight chapters deal only with organisation of statistical data and simple descriptive statistics and the last eight chapters cover simple tests of significance, including chi-square and binomial tests, and elementary inference and estimation procedures. Tables and diagrams are simple and clear, the text is unhurried and the many examples and exercises have been carefully chosen to illustrate some of the main uses of each statistic.

This concentration on specific computational and procedural details, however, results in the main discussion of probability concepts—two meagre pages—statistical inference and hypothesis testing being squeezed into three middle chapters and it is here that the main weakness of the text becomes most apparent. The authors apparently assume that mere exposure to some rules of use and computation of each statistic will simply and miraculously ensure the development of statistical wisdom. The logic of the rules and procedures is ignored or only cursorily discussed. Equally, some basic concepts, e.g., moments, expectation, degrees of freedom, random sampling, etc., are not only introduced and used without adequate discussion but are described in terms so concrete as to maximise the probability of the reader ultimately joining those reponsible for the present litter of statistical nonsense in psychology and education.

Typical is the treatment of 'expectation.' Never defined, it is merely illustrated in terms such as "if a coin were tossed forty times (i.e., a random sample of forty tosses), we would *expect* twenty heads." This reviewer, for one, would expect anything but twenty heads, but would define the mathematical expectations as twenty heads. Nowhere does the text distinguish between mathematical expectation, definable here as a parameter of a hypothetical sampling distribution, and the personal expectation of some possibly misguided psychologist.

As a somewhat restricted introduction to elementary statistics, then, the book has merit, particularly for those whose interests lie in psychological testing rather than in experimental design. Supported by adequate teaching to counteract its concrete approach, it could provide a useful classroom text. But, in this reviewer's opinion, its ability to fulfil the promise of its title and the avowed purpose of its authors is, even at best, only marginal. . A. B. Royse.

Dinkmeyer, Don. C. (1965). *Child Development—The Emerging Self.* New Jersey: Prentice-Hall, pp. xi+434, no price.

Another massive American production in a field so lacking in British equivalents that one cannot escape the fear of intellectual neo-colonialism!

How does this one differ from the others? First—layout. Instead of boxes to break up the text, we have other devices. Recommended reading is suggested in the body of the text instead of in footnotes and chapter endings, the educational implications of what has been said are listed at the end of key chapters, and questions appear at strategic points in each chapter. This last idea has possibilities, but the questions here are starting points for projects rather than immediate checks on comprehension.

Next—context. As often happens, the material is encyclopaedic, up-to-date, but unselected, so that major issues fail to stand out and the writer's own line tends to get lost in the detail. As opposed to a behaviourist—mechanistic interpretation of behaviour, Dinkmeyer is really interested in the unique pattern of interaction between internal dynamic forces and external influences in every child's life, and especially in the resultant, developing self concept and style of life with which an outsider must emphasise if he is to discover the ideographic laws of that child's behaviour. Here Dinkmeyer's thinking is ultimately derived from Adler, who is frequently quoted, and this holds also for his second main pre-occupation—that democratic procedures at home and at school are the best way of ensuring optimum personal development. Surprisingly the material quoted for this is untypically out-of-date, e.g., the early Lippitt and White experiments are repeatedly and uncritically quoted as evidence. Generally, however, the writer's views are interesting and persuasive, but so obscured by other detail that one is convinced that it is impossible simultaneously to compile an encyclopedia on a subject and also to present one's own views.

The organisation of material is orthodox. There is an excellent chapter on techniques of observation and recording, but the attempts to summarise all relevant theories in the chapters on motivation, intelligence, learning, etc., are inevitably condensed to the point of incomprehensibility. For example, Piaget's theories are squashed into two pages, and eleven different theories of motivation are summarised in just over three pages. The chapters on emotional, social and personality development are good, but more than usually inclined to spill over into each other, resulting in a lot of unacknowledged repetition, especially in a rather eccentric preoccupation with the influence of the family constellation.

Because it is a compromise between a general text and the exposition of a particular view of child development, I cannot see many people using this book other than as an up-to-date source book of references and American research findings. RALPH DUTCH.

EDWARDS ELWYN (1964). *Information Transmission*. London : Chapman and Hall, pp. 133, 15s.

This little paperback sets out to be exactly what its name suggests. It contains chapters on the classical equations for measuring the ' information content ' of a signal, and the rate of transmission of information through a human operator. *Information* here is the information of the engineer, that is, a function of the uncertainty of a signal, measured in such a way that halving the uncertainty in a situation gains one ' bit ' of information.

Information theory has had a tremendous impact over the last decade on experimental psychology, and while the first flush of enthusiasm has now paled to a more modest glow, much remains of value, and it is a field of which anyone wishing to be *au fait* with current trends should be aware Rather many psychologists find the mathematical side of the field difficult, and Mr. Edwards has tried to cover this by including a chapter on basic mathematics before deriving the information theory equations. While this is potentially of value, it still seems likely that those who find mathematics difficult, will find the pace of the book rather rapid, and those who find it easy will not need the introductory chapters.

A number of references are given to experiments in which information theory has been applied, and these are up to date with the date of publication There is a short chapter on networks of articial neurones which is interesting but not really germane to the main development of the book.

As a whole, it cannot be regarded as a substitute for Attneave's monograph of a few years ago, but should be found useful by students who need something to help them understand lectures in information theory. It is well produced with ample graphs, diagrams and tables for calculating information values. NEVILLE MORAY.

EVANS, K. M. (1965). *Attitudes and Interests in Education.* London : Routledge and Kegan Paul, pp. 168, 16s.

Dr. Evans' book " is intended as an introduction to some of the work on attitudes and interests which has been published since about 1930." There are six chapters on attitudes ; their development and modification, assessment, and organisation and (in greater detail) attitudes to authority, to the self, and intergroup attitudes. Two chapters deal specifically with interests ; their origins and development, and their assessment by questionnaires. The two final chapters discuss the relevance of attitudes and interests to achievement and to teachers and teaching.

Though said to be suitable for advanced courses in university institutes and departments of education, the text is commendably lucid and readable; and if the few references to factors were explained, it could be used with profit in ordinary courses. It would require extension by discussion, but might be made the basic text for a seminar whose members were expected to follow up, by allocation, the necessarily condensed items by using the comprehensive bibliography the text provides.

The skilful condensing of so much material merits remark, especially since the book is much more than a catalogue and contains many relevant evaluative comments, such as that factor analysis cannot produce information that has not been introduced into the test-material in the first place, or that much of what children read may be dependent to a large extent, not upon their free choice, but upon what books happen to be in their homes. Your reviewer would have welcomed an extension of comment to cover the whole area from an educational standpoint. When we claim to be assessing an ' attitude,' for instance, precisely what do we think we are doing ? Bluntly, is ' attitude ' itself a good idea ? But that is scarcely a fault, and Dr. Evans' text is warmly recommended to all who are engaged in teaching, and particularly to those who need a reasonably brief, comprehensive, readable, up-to-date account of an area of investigation in which they are not already expert. JAMES M. THYNE.

FOSS, B. M., (Edit.) (1965). *Determinants of Infant Behaviour III.* London : Methuen, pp. xiii+264, 50s.

This book contains an account of the third meeting of a group of psychologists, zoologists, and psychiatrists who have been coming together under the auspices of the Tavistock Clinic at two-yearly intervals in order to discuss research results and theoretical issues in the field of infant behaviour, with particular reference to the study of mother-infant interaction. Like its two predecessors, the present volume covers a considerable range of material, extending from " The basis of synchrony in the behavioural interaction between the mother and her offspring in the laboratory rat " to " The course of infant smiling in four child-rearing environments in Israel." Each contribution is followed by a judiciously edited version of the discussion which followed the presentation of the various papers.

One or two of the reports to be found here are excellent and merit careful reading. The paper by Rosenblatt on mother-young interaction in the rat is particularly worth mentioning in this connection, for it contains a most stimulating mixture of careful experimental work and provocative theoretical interpretations. Yet, as a whole, the volume is disappointing : amongst the reports of human studies in particular, the inclusion of methodologically and conceptually rather naive case studies depresses the tone of the whole book and leads one to rather sad conclusions about the advancement of knowledge among the clinically oriented workers in this field. Moreover, most of the systematic research papers included here are either merely preliminary reports or contain nothing but findings already well-known from other sources. This mixture of the trivial and the unexciting is, unfortunately, not improved by the reported discussions of the participants, for though they may have been highly stimulating to those present at the time on account of the sheer possibility of interchange, for the reader little benefit accrues from the rapid question-answer sequences that characterises most of them. There are some notable exceptions, such as is found at one point following a paper by Gewirtz on social learning under different child-rearing conditions, when a very high-level discussion takes place on the way in which computer methods may be used in the analysis of social behaviour sequences. By and large, however, this type of public eaves dropping does not make very rewarding reading.

It may well be, as the Foreword of the book puts it, that " within the group of participants a virus of co-operation is now strongly at work." One must hope, however, that future volumes in the series will contain rather more definitive and stimulating material, else the aim also expressed in the Foreword, that " the published proceedings infect others with the same virus," is unlikely to be accomplished.

H. R. SCHAFFER.

HORST, PAUL (1965). *Factor Analysis of Data Matrices.* New York: Holt, Rinehart and Winston, pp. xi+730, 120.

This is a very fine volume indeed. Professor Horst covers his chosen ground clearly and directly. His treatment reflects both his extensive research on factor analysis and prediction in industry and education and his considerable experience as a teacher.

To Horst, factor analysis is concerned basically with the parsimonious description of natural phenomena and with the reduction of the dimensionality of an empirically defined set of variables. He neatly sidesteps epistemological issues relating to the existence and nature of factors and emphasises the properties and computational procedures of the various solutions discussed.

The focal point of any factor analysis is the data matrix. Derived matrices cannot provide or generate information not manifest or latent in the data matrix itself. (Horst returns time and again to this theme.) This emphasis is, perhaps, the most important feature of the book. Too many writers have presented factor analysis merely as a method for analysing a correlation matrix. Horst restores the fundamental postulate to its proper position as *the* basic equation of factor analysis.

Seasoned veterans may find Horst's matrix algebra somewhat unorthodox. For example, the all-hallowed latent roots and vectors or eigenvalues and eigenvectors of a square symmetric matrix emerge naturally as a special case of the more general basic structure of a matrix and the problem of rank and the general rank reduction theorem are treated entirely without the aid of the traditional determinants. The introduction of these innovations and numerous other innovations in terminology is thoroughly vindicated by the use to which they are subsequently put.

Each of the methods chapters is organized as follows. An initial section describes the salient features of the procedures covered in the chapter. Computational equations, detailed computational steps, and numerical examples are presented for each procedure. Finally, mathematical proofs and derivations for each of the procedures are presented in an appendix at the end of the chapter. Thus, the book is readily adaptable for courses at elementary, intermediate or advanced levels of instruction.

A bonus feature is the appendix of computer programs for most of the procedures discussed. These are written in Fortran II and are obviously meant to be prototype versions rather than polished and flexible production versions. As such they are quite impressive and form a valuable supplement to the textbook material on the methods.

Much of the material presented is either new or has hitherto appeared only in relatively inaccessible project reports, in technical monographs, or in the proceedings of conferences on specialized psychometric topics nad thus becomes readily available for the first time in this formal textbook treatment.

This reviewer would have preferred greater emphasis on the problem of the number of factors and more discussion of rotation toward oblique simple structure as well as an explicit treatment of procedures for matching factors and for relating factors across studies. The latter methods are treated only tangentially in the chapter on multiple set factor analysis.

But make no mistake. This is an important book. It will become required reading for all psychometricians—master and novice alike—and will take its rightful place as a standard reference and as an authoritative textbook on the subject.

P. O. WHITE.

HEBRON, MIRIAM E. (1966). *Motivated Learning : A Developmental Study from Birth to the Senium.* London : Methuen and Co., Ltd., pp. 264, 35s.

To set out to report on all those phenomena within psychology which in one way or another are related to ' motivation and learning,' and to aim to report on motivation at all ages " from birth to the senium," represents a task worthy of Hercules. That Dr. Hebron has in large measure succeeded in this self-set task makes this volume worthy of note Furthermore, rather than present a mere catalogue of studies of motivation, she attempts some integration. Hebron suggests that consideration of the various current theories of motivation can lead to the development of " a profile of the various dimensions of motivation, which is characteristic of each developmental phase . . . " Motivation itself is regarded as a " global aspect of personality," having several major dimensions which are all continually operative. Against this background, the differences which exist between individuals in capacity to learn are regarded as differences in activities of the brain, prior to mental operations, and so learning and motivation are seen as being closedly integrated in a continuous process This conception dictates Hebron's further study, which considers learning and motivation together, and allows her to integrate data at each age-level so that typical developmental-stage profiles emerge.

This approach stressing integration of widely divergent studies naturally leads to some over-simplification. Statements appearing as evidenced fact might well have been better made as speculative comments. This is particuarly so in the background chapter on the physiology of motivation and learning. For example, no reference is given for the statement that " the extent to which sameness can be endured . . . can be increased by experience "—although the statement may well be true. Moreover, current thinking does not appear to confirm that " chemical changes in nerve endings " keep the brain in a state of rhythmic readiness, as Hebron suggests. The present reviewer's knowledge of contemporary neuro-physiology is too slight for him to make an independent judgment, but he is assured by a colleague of some standing in this field, that a more likely explanation of rhythmic readiness depends on the relative premeability of the neuron membrane. The danger of the somewhat naive phsyiology presented by Hebron is that it is used to interpret a wide range of motivated learning data, particularly for the early developmental stages ; such interpretations themselves must, therefore, suffer accordingly. The reviewer would stress, however, that such mis-statements are much less frequent in Hebron's account of non-physiological matters. Where they do occur, they can be accepted as small price to pay for the valuable addition to psychological literature of this extensive, yet critical review, which often succeeds in linking areas hitherto held to be separate.

It is, perhaps, this extensiveness and comprehensiveness which are the major assets of the book. It is noteworthy, too, for its coverage of British and Continental research, so often forgotten in current American texts. At the same time, the very comprehensiveness which gives the text its value, provides difficulty. So many facts, ideas, theoretical formulations and the like, are presented and related to the central theme, that the book has to be read very carefully if its major findings are to be grasped. The fact that for the present reviewer the style of the book hindered, rather than helped, this task was his misfortune The author's use of the adjective ' seminal ' at least five times in the first thirty pages, however, strikes this reviewer as being, like the book, pregnant with possibilities. G. SHOUKSMITH.

KELLMER-PRINGLE, M. L. (Edit.) (1965). *Investment in Children.* London : Longmans, Green and Co., pp. xi+180, 12s. 6d.

This book is based on the Proceedings of the first Conference held by the National Bureau for Co-operation in Child Care, whose aim is " to make more readily available existing knowledge on children's development and needs, and also to improve the lines of communication among all the disciplines professionally concerned with promoting healthy child growth." 'Prevention' was the theme of the Conference, and the essays in this present book are based on the papers read on that occasion. The place of education, paediatrics, social work and psychiatry in preventive work with children and their families is considered, and both the theoretical and administrative aspects are included Each discipline is represented by two contributions, one on theory and policy, the other on administration and practice The writers are well-known authorities in their respective fields. The final two chapters are by the editor, Dr Kellmer Pringle, who attempts to draw together the themes and suggestions contained in the various contributions

The book represents not so much an attempt to suggest remedies, as an attempt to pose questions, raise issues and give consideration to various points of view. It contains much worthwhile information—Chapter 3, for example, gives a very clear outline of the present state of knowledge about mentally handicapped children. The book is of value in providing within a single volume contributions from representatives of the various disciplines concerned in promoting the welfare of children. Each contribution, though interesting in its own way does not contribute to a unified picture of the field, and the reader is left rather confused and dissatisfied by the present set-up. One cannot, however, criticise the contributors in as far as they are merely relating the present rather confusing situation which *does* exist in welfare work in general. Possibly the contributions have lost something in being transferred from 'live' papers to written essays, in spite of the fact that they were amended before publication. Even with its limitations, the book does repay study by those concerned with the welfare of children, whether the so-called 'normal,' maladjusted or delinquent. MARGARET M. CLARK.

KOSKENNIEMI MATTI (1965). *The Development of Young Elementary School Teachers.* Helsinki: The Finnish Academy of Science and Letters, pp. 635, no price.

In many ways, this is a remarkable book. It is the report of an entrance examination experiment in one of the training centres for primary school teachers in Southern Finland. The study was a response to the State committee's need to prepare legislation to expand and modify teacher education—a case of educational research preceding educational decision.

It describes an attempt to devise an entrance examination which would preferentially select potentially good teachers. Thus, of the 405 original applicants (160 men) for the seventy-two places in one college which trained primary school teachers, 175 candidates (110 men) were given the entrance tests to select the forty-eight men and twenty-four women students to be admitted. The successful applicants were followed subsequently through their courses and into the profession. A broad spectrum of academic, psychological and sociological data was obtained, including estimates of satisfaction and success in the teaching situation. A detailed examination was then undertaken of a large number of factors which it was thought could possibly impinge on the development of the teacher, and the treatise reported these attempts to get to the bottom of what became obviously a multi-dimensional problem. Its findings were numerous rather than definitive, as intra-group studies ended up with very small sub-groups indeed In fact, as a study its major weakness was its limited subject group (both in number and type).

As a report it seemed to be more a collection of chapters and sections than a presentation of a single complete picture. This criticism is perhaps harsh, as the complexity of the problem tackled made straightforward treatment impossible and an acceptable, compact answer unobtainable, particularly in the light of the size of the sample. Even so, it did not make for light reading

The results included both group analyses and case histories, and, although it admits to posing more questions than it answers, the report has many points of interest for any one working in the field of teacher training. Certainly, a copy of it should be on the library bookshelf of every department which has a research interest in this field. K. B. START.

MACNAMARA, JOHN (1966). *Bilingualism and Primary Education : A Study of Irish Experience.* Edinburgh : University Press, pp. viii+173, 35s.

Can the answers to urgent political and educational problems, the requirements of a doctoral dissertation, and the ingredients of a readable book, all be supplied by the same piece of research ? Father Macnamara has made a valiant attempt to prove that they can, but his project was too ambitious to succeed. Much time and planning are required to carry out the first survey of attainments in a country where testing is a novel procedure, even when others are willing to do the necessary field work. Unfortunately, the inadequacies of this investigation preclude it from being used as evidence in the controversy over teaching in Gaelic, and the author should be thankful that it was accepted for a Ph.D. As to the book, he might find consolation in a recent remark by Mace in the *B.J.P.* that : " The traditions and conventions of the universities make theses as presented for higher degrees almost invariably unpublished."

The review of the literature on bilingualism is excellent, and there is a fascinating, but all too brief, glimpse of the Irish educational scene. The fifth standard, for instance, which was sampled in this survey, has an age range of 8 to 16 years. Without previous experience of tests the sample took five in one day, some of which were specially prepared, with norms derived by extrapolation from British conversion tables. To obtain comparable English and Irish versions of Schonell's Arithmetic Tests, a preliminary alternation was carried out with a small group of Dublin children. The boys taking the English-Irish sequence showed an improvement of 9 points on the second occasion, while the Irish-English procedure yielded 161 extra points. This difference is found not significant by analysis of variance, which was used in spite of lack of homogeneity in the groups. Later, in excusing a similar use of this technique, the author says that . " The effect of heterogeneity of regression on the F test, like the effect of heterogeneity of variance, is probably slight, tending to make a mean difference appear significant when, in fact, it is not." (p. 104).

Had Dr. Macnamara confined himself to the construction and standardisation of tests for Irish children of all ages, he would have produced a worthy thesis and done a great service to educationists in Ireland. He would also be a step nearer to carrying out the research, which is still badly needed, on the effects of bilingualism on Irish education, and hence to the book which one day he will undoubtedly write.

G. W. H. LEYTHAM.

MUSGROVE, F. (1966). *The Family, Education and Society.* London : Routledge and Kegan Paul, pp. viii+156, 21s.

Dr. Musgrove argues that the family, while it has lost most of its *powers*, today enjoys perhaps a greater degree of *influence* than it has ever had before. Far from being in decay, it is our most successful and flourishing social institution. There is no real evidence, he claims, that marriage is less stable than in the past, and many indications that the influence of parents on the values, standards and behaviour of their children is by far the most important influence (indeed, Dr. Musgrove thinks it may be too important). Research, even in the United States, does not support the view that the standards of parents are, in important respects, rejected in favour of those of ' peer groups .' There had been, Dr. Musgrove claims, a strong tendency, reinforced by such technical development as television and cars, for families to become more self-contained and for their members correspondingly to be less involved in social activities outside the home. Dr. Musgrove sees this as " a serious threat to social cohesion," and one related to the dangers of excessive assortative mating, which tends to have a socially divisive effect at a rather different level.

He also sees a threat in the argument that education should be tailored to meet the demands of *parents* as consumers. " Schools which are extensions of families narrow the child's social range. They stunt his intellectual development and blunt his moral awareness and sympathies . . . If schools are to become our principal means of social integration—even our principal means of education—they must regain their truly public character."

In a chapter entitled " The Good Home," Dr. Musgrove makes out a good case for the view that the ' home treatment ' most likely to lead to ' success ' is *demanding* treatment. Indeed, he remarks on p 86, " Of particular importance are the child's relationships with its parents. If these are close, warm and affectionate, he is likely to be handicapped for life." He suggests that the idea that a boy needs a strong, successful father with whom to ' identify ' is simply wrong, and that the boy is likely to do better if he has " an inadequate and feckless (if ' demanding ') father."

Chapter 6 in Dr. Musgrove's book reports an inquiry into the nature and degree of ' satisfactions ' (for the child) provided by the home, the club, work and the school. The conclusions in general are that home, work and (to a lesser extent) the club, *do* tend to meet young people's expectations to a considerable degree, but that the grammar school, in particular, may be failing in its task.

The conclusions at which Dr. Musgrove arrives are that the school should offer its pupils at least some of the things which are not and cannot be provided within the family, that the school should act as a bridge between the self-contained family and society and that in order for this to be done effectively, *all* secondary schools should be comprehensive and public in the proper sense of the word. The reviewer has a lot of sympathy for Dr. Musgrove's point of view. However, in his efforts to state a good case, Dr. Musgrove perhaps tends to overstate it, and to make statements of opinion as if they were statements of fact, or at any rate as if they were more universally accepted than they are. On p. 135, for example, he says, " It is the business of education in our social democracy to *eliminate* the influence of parents on the life-chances of the young." Such a statement surely cannot be accepted unreservedly. Again, on p. 136, he states that " the aim and ideal of the meritocracy is purely chance or random association between the status of parents and the eventual adult status of their children," a statement which would appear to rest on questionable assumptions.

To sum up, this is a provocatively written book, the arguments and conclusions of which, even though they will by no means be accepted by everyone, deserve careful attention and consideration. They are concerned with things which really matter.

D. GRAHAM.

NEUBAUER, P. B. (Edit.) (1965). *Concepts of Development in Early Childhood Education.* Springfield, Ill. : Charles C. Thomas, pp. ix+149, $6.50.

It is difficult to understand why this volume was published at all. It represents the proceedings of a conference held by the Child Development Center of New York, this being a psychiatric clinic for pre-school children with a therapeutically oriented nursery school attached to it. The intention was to bring together educators, clinicians, and research workers from the fields of experimental psychology, psychoanalysis, education and anthropology, and attempt to find bridges not only between the diverse attempts of these various disciplines to understand the process of development but also, and primarily, between theory and practice.

The aims are admirable, but the volume falls far short of achieving them. This is by no means due to the individual participants, whose contributions are often of a high intellectual order. It is rather that the book as a whole is very much lacking in cohesion, fails to develop any particular theme, and, above all, was much too ambitious an undertaking in view of the nature, limited number, and considerable diversity of the topics discussed. There are four contributions dealing respectively with the concept of developmental stage, the role of language in early childhood, the contribution of nursery school experience to personality development, and the subject of death fear in young children's play. The best of these is the first—inevitably structured around Piagetian theory and occasionally straying, somewhat vaguely and timorously, into attempts to spell out the consequences of these concepts, originally derived from intellectual development, for emotional growth. The chapter on language is written from a psychoanalytic point of view and fails to make any reference to the considerable body of data on this subject contributed by anyone outside psychoanalysis ; that on the developmental contribution of nursery school attendance is so vague and general that it is difficult to understand the point it is attempting to make ; and the chapter on death fears (contributed by an anthropologist), while containing some fascinating material, occupies very much of an isolated position in the book as a whole.

In his summing up the Chairman of the conference stresses how readily the discussions following the papers moved from concept to practice, thus proving the value of having both theoreticians and practicians around the same table. It would be interesting to ascertain how much, in fact, the latter benefited from the former and whether educational methods were changed one iota as a result of this conference. The gap between theory and practice is still a wide one, and it is unlikely that this book can make a substantial contribution to its bridging.

H R. SCHAFFER.

SCHERER, GEORGE A. C., and WERTHEIMER, MICHAEL (1964). *A Psycholinguistic Experiment in Foreign-language Teaching.* New York : McGraw-Hill, pp. 256, 60s.

This slim, but expensive book, should be of interest to foreign-language teachers and, in parts, to psychologists interested in language. The authors, a language professor and a psychology professor at the University of Colorado, state that their work had two main purposes, to compare audiolingual and traditional methods of teaching German, and to devise tests to measure a variety of aspects of second language learning.

Students beginning German were randomly assigned to either audiolingual or traditional methods of teaching, and the experimenters were able to show that the two groups of students were comparable on a variety of control measures. The instructors were also matched. The teaching methods differed in a number of respects, the most pervasive being that the traditional method was weighted more towards reading and writing, the audio-lingual method more towards listening and speaking.

The authors developed their own test materials for listening, speaking, reading, writing, translating, direct association (or, ' thinking in German ') and various aspects of motivation and attitude. The size and pattern of intercorrelations amongst the tests suggest reasonable success with regard to reliabilities and validities. A rough and ready statement of results would be that students taught by the traditional method were better on tests of reading, writing and translation ; students taught audiolingually were better on tests of listening, speaking and direct association ; overall it is difficult to say that one method was better than the other.

The book is a solid report of an ambitious study. The pre-test and post-test materials, the teaching materials and procedures, and the results themselves are all presented in considerable detail. Indeed, the book is interesting as a case study of the problems involved in attempting to conduct a carefully controlled experimental study in a realistic teaching situation.

For foreign-language teachers, the moral seems to be that if you want students to learn to read and write a foreign language, teach them to read and write it ; if you want them to speak the language, teach them to speak it. For psycholinguists, the moral must be that written language is not just a logical extension of spoken language, and the different patterns of results found with tests of language production on the one hand, and language comprehension on the other, provide food for thought too.

COLIN FRASER.

SLAUGHTER, S. S. (1966). *The Educable Mentally Retarded Child and His Teacher.* Philadelphia : F. A. Davis, Co., pp. 191, $4.50.

The author of this book, Mrs. Stella S. Slaughter, is well-known in American educational circles and has for many years interested herself in the mentally retarded child. The present volume is " a textbook for teachers and prospective teachers " and describes the schooling needs of children of I.Q. 50-70 approximately. Every aspect of class-room work is covered including advice on planning the timetable and on the teaching of the various basic subjects. Each chapter concludes with a few suggestions for further reading, with some indication of the ground covered by each title listed. Chapters on the principles of teaching the mentally retarded, their general characteristics and needs and the qualities required in their teachers are all considered.

The book is a straightforward and very comprehensive statement of Mrs. Slaughter's extremely sensible views. The arrangement is clear · each chapter treats a distinct aspect of the subject and each paragraph starts with a sentence in heavier type announcing its theme. Every conceivable practical issue seems to be dealt with. In so far as there is emphasis, it is on preparing the child for ultimate adjustment in the community. This will involve the teacher in work and contacts outside the class-room, and Mrs. Slaughter devotes chapters to the consideration to these matters. A few well-chosen case histories illustrate certain points.

As a straight-forward textbook introducing the student to a particular educational field, the book has some merits and should win a place in the libraries of colleges of education. It is easy to read, thorough, and practical. As the embodiment of the views of an experienced and wise guide, it is convincing and deserves attention. But it is somehow unexciting. It spells out everything and never quite gets off the ground. O. C. SAMPSON.

THOMAS, A., BIRCH, H. G., CHESS, STELLA, HERTZIG, MARGARET E., and KORN, SAM. (1966). *Behavioural Individuality in Early Childhood.* New York: New York University Press, and London: University of London Press, pp. xii+135, 21s.

Since 1956, the authors have been making a longitudinal study of 130 children. This is a report of their methods and findings in relation to the first two years of life, preceded by a discussion of the theoretical considerations which led to their choice of variables to be studied.

Believing that no fully adequate theory of development is yet available, they chose " an inductive hypothesis-producing rather than a deductive hypothesis-testing " approach. Considering also that there has been in the past too much emphasis on environmental determination of personality, they emphasise the importance of studying individual differences in the organismic substrate on which these environmental factors act, as a preliminary to exploring the interactions of these two. They are, therefore, concerned with describing behaviour rather than interpreting it, and with finding categories of behaviour within which individual difference could be seen.

By itself, this approach would seem to leave unanswered many of the most important questions. " The statement, ' He cried because he was hungry ' was scored only in relation to crying. The child might or might not be hungry, but there is no doubt that he did cry." But what every parent wants to know is *why* baby is crying—not from pre-occupation with any particular theory, but because the success of any practical efforts to change the behaviour depends on its correct interpretation. Similar questions arise for other educators and clinicians.

Presently available interpretive systems, particularly the psychoantlytic, are demolished by the authors, who offer at present nothing in their place. But they do show themselves to be fully aware of the complementary approaches needed, and make clear that the present study is but the first stage. And if one feels that their criticisms of psychodynamic theory are less than fair, one cannot but agree with many of their criticisms of experimental methodology based on it.

The present report analyses the observed behaviour into nine categories: activity, rhythmicity, adaptability, approach, threshold, intensity, mood, distractability, persistance. Early profiles over these categories were found stable (in the same family environment) over the first two years of life: which is to say that characteristics of behaviour identifiable during the first six months were good predictors over the following year-and-a-half.

The last chapter discusses some of the practical implications of these initial and persistent individual differences, in terms of the differing reactions to be expected from the same environmental factors. A beginning is thus made with the next development, of studying individual differences of interaction: and I await further instalments with much interest RICHARD R. SKEMP.

TORRANCE, E. P. (1965). *Constructive Behavior: Stress.* Belmont, California: Wadsworth Publishing Co., Inc., pp. x+432, 56s.

Since the early 1950's, E. Paul Torrance and his colleagues have published a stream of research reports, first on adult, often military, behaviour under conditions of extreme stress or danger of survival, and then on thinking processes, especially creative thinking, with emphasis on the school child as a subject of study. In this book he now combines findings from these apparently dissimilar lines of research into a comprehensive theory of the psychological environment which encourages constructive behaviour in the face of stress and hence favours mental health.

The book must be evaluated in three aspects. First, as a review of the literature, we are presented with a refreshing slant on mental health, personality development and the interactions of stress reactions and cognitive processes. (It should be noted that I give the term ' cognitive ' a much wider meaning than the author—following Guilford—would allow, and include the creative processes of divergent production and evaluation). There is, therefore, a valuable summary of research but doubts are felt over the citation of different types of experiment, laboratory and in the field, with equal enthusiasm and equal confidence in the experimenters' conclusions. There is even a distressing tendency to give weight to anecdotal evidence and studies are quoted more as illustrative examples than as necessary support for a theory.

Secondly, the author aims to *teach* the reader to identify problems of mental health as they develop and to cope with them in group members with whom he may work. There are exercises set for the reader and we are provided with lists of mental health danger signals (" failure in studies, athletics, friendships, dating . . . excessive outside activities . . weak realization of reality . . . unusual docility ") and with guides to improving memory (" File (what is to be remembered) in its proper place in his memory system ") and other cognitive functions. I found much of this irritating for its lack of precise definition and even of meaning in practical situations. Nevertheless, there is practical guidance amongst the sections of advice. In particular, organizers of groups, whether social, military or educational, could benefit from recognition of the stigmata of healthy and unhealthy groups discussed here, and educationists, too, will find direct guidance which, though intended for the United States is applicable also in Britain.

But it is thirdly, as a presentation of a theory that the book must finally be judged. It must be admitted that the theory is weak in its uncritical reliance on ' creativity ' research, which has been attacked by the statisticians, and on personality concepts of the unconscious and even the group unconscious. And by assuming that all life stresses can be ordered on continua of severity and duration, the implication that no other stress variable is relevant, arises but is not defended. However, even weak theories can be useful or can be made more useful, and this book will have the resounding merit of stimulating research. ANNE BROADHURST.

VERNON, JACK (1965). *Inside the Black Room.* London: Souvenir Press, pp. xvii+203, 21s.

The ' black room ' of the title is the room in Princeton in which Dr. Jack Vernon's subjects underwent ' sensory deprivation ' for varying periods up to ninety-six hours. The subjects were some 100 male paid volunteers (" 20 dollars per confinement day or any partical day ") ; they were mainly graduate students from various departments in Princeton, between 20 and 32 years of age. The room was lightproof and soundproof. Subjects were provided with a bed, with lavatory facilities, with ample sandwiches, fruit, soup and water ' in a picnic ice box ' and with a 'panic button' which they could press at any time they wished to end the experiment. Once having pressed the button, the ' failures ' were let out and took the various tests which were given to the ' successful ' 96-hour subjects but, even if they wished to do so, they were not allowed to re-enter the experimental room.

The dedication, " To three fine people, B.J.V., S.M.V., V.L.V." gives the reader an excellent idea of what to expect from the book itself. It is an unsophisticated, honest but rather childlike effort, written in an excessively popular style, evidently intended for the uninformed layman. Explanations are given, for instance, of such terms as ' reaction time,' ' decibel ' and ' control subjects. There is no index.

Where the numbers of subjects are given (and they often are not) for the various experiments, these are sometimes woefully inadequate. For example, in the chapter on " The effects of sensory deprivation upon learning," the sub-groups are on more than one occasion as small as n=4. Yet paucity of numbers is apparently never considered as the reason for inconclusive results.

In fact, experimental naiveté is the keynote for a good deal of the book. Let us give three examples, taken at random : " The total word count of the stories was used as a measure of the ability to maintain thought." " They may have repressed certain items, but that doesn't seem likely, for they spoke quite freely when reminded." " We may conclude that the effects of sensory deprivation are similar to those of brainwashing." The last statement is particularly striking when it is recalled that the Princeton subjects were volunteers, knew that the experiment was intended to last for ninety-six hours, were aware of the presence of an observer outside the room and had a ' panic button ' available.

The book may be of interest, however, to readers hitherto unfamiliar with work on sensory and perceptual deprivation. It covers a great many topics, such as dreams, hallucinations, thinking, learning, the estimation of time, pain perception and loss of weight. It gives evidence of the enormous individual differences found in sensory deprivation experiments, both in tolerance and in effects. In the reviewer's opinion, it resembles other work in this field in its apparent underestimation of the role of *social*, as opposed to sensory deprivation. A. W. Heim.

Wallach, M. A., and Kogan, N. (1965). *Modes of Thinking in Young Children.* New York : Holt, Rinehart and Winston, Inc., pp. viii+357, $8.00.

Creativity, a popular modern topic, and the distinction between creativity and intelligence, are the subjects discussed in this latest book from the pens of Wallach and Kogan. The importance of this volume, and it is an important book, lies in the treatment offered the two concepts, which are regarded not as two areas of ability but rather as two modes of thinking which are functionally related to a wide range of behaviours. A review of the experimental literature, particularly the work of Torrance and his associates, and the work of Getzels and Jackson, suggested to the authors that on evidence thus far provided no distinction could be made between intelligence and a second aspect of cognitive functioning, labelled ' creativity." A further review, of introspective as opposed to psychometric accounts, however, showed that the majority stressed associative freedom as the central aspect of creativity and regarded playful situations and permissive task attitudes as most favourable for creative production. Thus, an associative denifition of creativity was adopted, a subject's creativity level being measured by his ability to generate unique and plentiful associates, in a generally task-appropriate manner and in a relatively playful context.

The neat procedure adopted by Wallach and Kogan was to take overall ' intelligence ' and ' creativity ' scores and then divide the total sample of 151, 5th grade pupils at the median on each factor so producing four sub-groups : the high creativity—high intelligence group, the high creativity—low intelligence group, the low creativity—high intelligence group and the low creativity—low intelligence group. The remainder of the work is concered with investigating the psychological characteristics of each group, comparing them on behaviour in the classroom, methods of categorizing and conceptualizing, sensitivity to physiognomic properties and degree of anxiety and defensiveness. A host of interesting findings are presented and both their theoretical and practical implications discussed in full ; some of these findings not only add to our knowledge, but call into question well-worn older conceptions. Educators, for example, will find a challenge in Wallach and

Kogan's conclusions that inferential abstraction, typically regarded as the most mature form of cognitive functioning, is not so important as the relative balance between various modes of thinking, including previously despised relational and thematizing modes.

As one would expect from Wallach and Kogan, the presentation and the matter are both good. One or two omissions, however, surprised the present reviewer. In the authors' discussion of the relation of anxiety to intellectual functioning no mention is made of the work of Eysenck and no consideration is given to non-linear relationships. In other review sections, too, there is a ' schools atmosphere, some authors cropping up time and time again whilst others working in the same area are excluded or mentioned only in passing ; in the section on conceptual style, for example, Witkin rates only a footnote, whilst in all sections, to the British eye, the work is noticeably trans-atlantic ! The generally excellent nature of the book makes one forget that the whole work stands or falls on the adequacy of the definition of creativity as a free, associative process. Again, the review presented, largely of anecdotal evidence, and again with important omissions, is not sufficient to convince the reviewer that this is the only possible interpretation of ' creativity ' Nor do the ' assessment games ' for measuring it seem anything more than ' tests ' given under permissive conditions with ' playful,' tension avoidant instructions. Whether or not it is creativity in full which is being measured, however, may well be immaterial. Wallach and Kogan, in their own words, " have isolated two fundamental, orthogonal modes of thinking, whose joint effects are felt in many areas of psychological functioning." In so doing, they have also initiated a much needed rapprochement between a concern for individual differences and parametric examinations of tasks and functions, which may well redirect future theory and research on children's thinking.

G. SHOUKSMITH.

PUBLICATIONS RECEIVED

A. G. Nikelly (1966). *Mental Health for Students.* U.S.A.: Charles C. Thomas, pp. xv+208, $7.50.

G. Vaughan (1966). *Psychiatry for Students.* London: George Allen and Unwin, pp. 276, 21s.

V. M. Axlind (1966). *Dibs: In Search of Self.* London: Victor Gollancz, Ltd., pp. 186, 25s.

R. R. Sears, L. Rau and R. Alpert (1966). *Identification and Child Rearing.* London: Tavistock Publications, Ltd., pp. xiii+383, 63s.

A. M. Clarke and A. D. B. Clarke (1966). *Mental Deficiency: The Changing Outlook.* London: Methuen and Co., pp. xxi+596, 63s.

G. A. Nuthall and P. J. Lawrence (1965). *Thinking in the Classroom.* New Zealand: New Zealand Council for Educational Research, pp. 60, no price.

D. J. Vail (1966). *Dehumanization and the Institutional Career.* U.S.A.: Charles C. Thomas, pp. xxvii+266, $10.75.

M. H. Sherman (1966). *Psychoanalysis in America: Historical Perspectives.* U.S.A.: Charles C. Thomas, pp. xii+518, $15.75.

M. M. Craig and S. J. Glick (1966). *A Manual of Procedures for Application of the Glueck Prediction Table.* London: University of London Press, pp. 44, 7s. 6d.

G. Mialaret (1966). *The Psychology of the Use of Audio-Visual Aids in Primary Education.* London: George G. Harrap and Co., pp. 225, 30s.

Hans Raj Bhatia (1966). *A Textbook of Educational Psychology.* London: Asia Publishing House, pp. vii+610, 60s.

M. Branch and A. Cash (1966). *Gifted Children.* London: Souvenir Press, Ltd., pp. 222, 30s.

N. Whalley (1966). *A Guide to the Preparation of Teaching Programmes.* Bristol: Teaching Programmes, Ltd., pp. vii+36, 7s. 6d.

P. I. Jacobs, M. H. Maier, and L. M. Stolurow (1966). *A Guide to Evaluating Self-Instructional Programs.* New York: Holt, Rinehart and Winston, pp. 84, no price.

L. Hudson (1966). *Contrary Imaginations: A Psychological Study of the English Schoolboy.* London: Methuen and Co., pp. 181, 25s.

O. G. Brim, Junr., and Stanton Wheeler (1966). *Socialization After Childhood.* New York: John Wiley, pp. vi+116, 38s.

Jean Drevillon (1966). *L'Orientation Scolaire Et Professionnelle.* Paris: Presses Universitaires De France, pp. 153, no price.

Education in Europe (1966). *Structure of University Staff.* Strasbourg: Council for Cultural Co-operation of the Council of Europe, pp. 148, 8s. 6d.

Education in Europe (1966). *The Teaching of Chemistry at University Level.* Strasbourg: Council for Cultural Co-operation of the Council of Europe.

P. Blommers and E. F. Lindquist (1966). *Elementary Statistical Methods in Psychology and Education.* London: University of London Press, Ltd., pp. xv+528, 50s.

P. Blommers and E. F. Lindquist (1965). *Elementary Statistical Methods in Psychology and Education.* London: University of London Press, Ltd., pp. 247, 17s. 6d.

M. Brearley and E. Hitchfield (1966). *A Teacher's Guide to Reading Piaget.* London: Routledge and Kegan Paul, pp. xii+171, 18s.

R. C. Craig (1966). *The Psychology of Learning in the Classroom.* New York: The MacMillan Co., pp. xvii+85, 10s. 6d.

G. L. Bond and E. Bond Wagner (1966). *Teaching the Child to Read.* New York: The MacMillan Co., pp. viii+404, 55s.

Allan Fromme (1966). *The Ability to Love.* London: Geroge Allen and Unwin, Ltd., pp. x+366, 30s.

A. G. Wedderspoon (1966). *Religious Education, 1944-1984.* London: George Allen and Unwin, Ltd., pp. 238, 21s. net, cloth, 12s. 6d. net, paper.

F. H. Hilliard, D. Lee, G. Rupp and W. R. Niblett (1966). *Christianity in Education.* London: George Allen and Unwin, Ltd., pp. 111, 18s.

Rivka R. Eifermann (1966). *Scripta Hierosolymitana, Volume XIV: Studies in Psychology.* London: Magnes Press, Oxford University Press, pp. 230, 43s. net.

G. Boas (1966). *The Cult of Childhood.* London: The Warburg Institute, University of London, pp. 106, 30s.

G. H. Hildreth (1966). *Introduction to the Gifted.* New York: McGraw-Hill Book Co., pp. xvii+572, 64s.

B. M. Foss (1966). *New Horizons in Psychology.* Penguins Books, Ltd., pp. 447, 7s. 6d.

Per Saugstad (1966). *An Inquiry into the Foundations of Psychology.* Oslo: Universitetsforlaget, pp. 103, 18s. 6d.

Jean Vial (1966). *Classes De Transition Et Classes Pratiques.* Paris: Presses Universitaires De France, pp. 135, no price.

O. K. Kyöstiö (1966). *Work Tests as Indicatorsof Vocational Fitness.* Helsinki: Suomalainen Tiedeakatemia, pp. 62, no price.

A. Lehtovaara and P. Saarinen (1966). *School-Age Reading Interests.* Helsinki: Suomalainen Tiedeakatemia, pp. 216, no price.

J. MacNamara (1966). *Bilingualism and Primary Education.* Edinburgh: University Press, pp. 173, 35s.

S. R. Slavson (1966). *The Fields of Group Psychotherapy.* New York: J. Wiley, pp. xiii+338, 15s.

INDEX TO VOLUME XXXVI, 1966

Articles :

Abilities, Secondary Education, Social Class and the Development of: W. B. Dockrell .. 7
Achievement, What Can Personality and Motivation Source Trait Add to the Prediction of School ? R. B. Cattell, A. P. Sealey and A. B. Sweney .. 280
Adolescent Girls, The Perception of Peer Personality: A. Morrison and H. J. Hallworth .. 241
Adolescents, A Comparative Study of Value Judgments of: H. J. Hallworth and G. Waite .. 202
Attainments and Abilities, Towards a Rationale for Testing in Africa: S. H. Irvine 24
Children's Behaviour, A Comparison of at Home and at School: S. Mitchell and M. Shepherd .. 248
Co-educational and Single-Sex Grammar Schools, Pupil-Teacher Relationships in: R. R. Dale .. 267
Co-educational and Single-Sex Grammar Schools, the Happiness of Pupils in: R. R. Dale .. 39
Conservation of Quantity, A Training Procedure for Acquisition of Piaget's: I. E. Sigel, A. Roeper and F. H. Hooper .. 301
Creativity and Intelligence: A. J. Cropley .. 259
Desirability Response Set in Children: S. B. G. Eysenck, I. A. Syed and H. J. Eysenck .. 87
English, The Validity and Reliability of Tests of Spoken: P. J. Hitchman .. 15
Errors, The Role of, in Learning with Feedback: W. B. Elley .. 296
Forced-Choice and Likert-Type Tests of Motivation to Academic Achievment, Evaluation of: C. E. Buxton .. 192
Guidance, a Programme of Educational for The Primary School: N. France and S. Wiseman .. 210
Intelligence Tests, the Relationship between Level and Speed on Two: A. J. Yates 166
Intelligence, the Relationship between and Anxiety, Neuroticism and Extraversion: P. Ley, M. S. Spelman, A. D. M. Davies and S. Riley .. 185
Intelligence Performance, Level Speed and Personality Factors of Young Children: A. J. Yates .. 312
Judgment of Adolescent Pupils, A Study of the Differences in: E. A. Peel .. 77
Lecture Time, the Effect of, on Learning: P. J. Holloway .. 255
Mathematics Test, A Study of the Effects of Instruction in a Multiple-Choice: P. H. Taylor .. 1
Personality Characteristics, the Relationship between and Scholastic Success in Eleven-year-old Children: J. Rushton .. 178
Personality Measures and their Relation to Teaching Ability: K. B. Start .. 158

Scottish Children, Differences in Test Score in Gaining Selective Places of, compared with those in England and Wales: J. W. B. Douglas, J. M. Ross, S. M. M. Maxwell and D. A. Walker .. 150
Social Needs and Satisfactions of Some Young People, Part I. At Home, in Youth Clubs and at Work: F. Musgrove .. 61
  Part II. At School: F. Musgrove .. 137
Summer and Autumn-born Children, A Comparison of the Performance of at Eleven and Sixteen: H. G. Armstrong .. 72
Teachers' Assessments of Primary School Pupils, Social and Educational Variables Relating to: D. McIntyre, A. Morrison and J. Sutherland .. 272
Verbalisation Effects in a Non-Verbal Intelligence Test: J. Radford .. 33
Vocational Suitability, The Prediction of, from Secondary Modern School Record Cards: J. H. Eyre .. 48

Research Notes :

Adult Education, A Null Relationship between Ecology and: T. Lee .. 100
Extraversion, Neuroticism and Academic Performance among Ghanaian University Students: P. Kline .. 92
Impaired hearing, the Social and Emotional adjustment of Children with: B. Fisher 319
Intelligence, A Note on Season of Birth and: A. B. Ojha, R. P. Kelvin and C. J. Lucas .. 94
Learning, The Application of Factor Analysis to Problems of: G. E. R. Burroughs 227
Personality Factors and Academic Attainment in Junior School Children: R. D. Savage .. 91
Puberty, the Secular Trend in, Has Stability been Achieved ?: P. K. Poppleton and P. E. Brown .. 95
Reader Course, An Investigation of the Craig: K. D. Duncan and R. Foster .. 327
Retardation, the Assessment of in Reading and Arithmetic in the School: R. D. Savage and D. J. O'Connor .. 317
Weight of an examination paper, What is the: J. M. Thyne .. 321

Critical Notices:
Concept Growth and the Education of the Child, by J. G. Wallace: E. A. Lunzer .. 124
Contrary Imaginations, by Liam Hudson: F. V. Smith 334

Thesis Abstracts:
Anxiety, the Effects of on Learning, Task Performance and Level of Aspiration in Secondary Modern School Children: U. Maguire 109
Auditory High Frequency Weakness, the Association of with Verbal and Written Comprehension and Expression: G. F. Reed 118
English, An Investigation into Two Methods of Teaching in Modern Schools: D. Cooper 105
English Language, A Comparison of the Errors made by Grammar and Technical School Pupils in Alternative Papers in G.C.E.: D. M. Brannan 121
English, Teachers' Assessments of the Aims of Teaching, in Secondary Schools: B. M. Casey 106
Mental Ability, Age and: Re-Testing with Thirty-Three Years' Interval: R. B. Burns 116
Reading Ability, the Relationship between, and Verbal Intelligence Test Performance: C. S. Riley 117
Science, Teaching of Elementary by Television and Other Methods: H. Barrington 114
Science Teaching, the Opinions of Science Teachers on the Objectives of: G. J. Pollock 112
Sociometric Comparison, A, of Well-Behaved and Badly-Behaved Classes: J. I. Rose 331
Thematic Apperception Test, Junior Children's Responses to the Murray and the Cattell Personality Questionnaire: J. M. Mundy 103
Town and Country, An Investigation into the Differences between Children in the Formation of Concepts of: J. F. Marsh 333
Transfer, The Effect of on the Level of Children's Adjustment to School: W. F. Murdoch 332

Authors:
Armstrong, H. G.: A Comparison of the Performance of Summer and Autumn-born Children at Eleven and Sixteen 72
Barrington, H.: The Teaching of Elementary Science by Television and Other Methods 114
Brannan, D. M.: A Comparison of the Errors Made by Grammar and Technical School Pupils in Alternative Papers in G.C.E. English Language 121
Burns, R. B.: Age and Mental Ability. Re-Testing with Thirty-three Years' Interval 116
Burroughs, G. E. R.: The Application of Factor Analysis to Problems of Learning 227
Buxton, C. E.: Evaluation of Forced-Choice and Likert-Type Tests of Motivation to Academic Achievement 192
Casey, B. M.: Teachers' Assessments of the Aims of Teaching English in Secondary Schools 106
Cattell, R. B., Sealy, A. P., and Sweney, A. B.: What Can Personality and Motivation Source Trait-Measurements Add to the Prediction of School Achievement? 280
Cooper, D.: An Investigation into Two Methods of Teaching English in Modern Schools 105
Cropley, A. J.: Creativity and Intelligence 259
Dale, R. R.: The Happiness of Pupils in Co-educational and Single-sex Grammar-Schools: A Comparative Assessment 39
Dale, R. R.: Pupil-teacher Relationships in Co-educational and Single-sex Grammar Schools 267
Dockrell, W. B.: Secondary Education, Social Class and the Development of Abilities 7
Douglas, J. W. B., Ross, J. M., Maxwell, S. M. M., and Walker, D. A.: Differences in Test Score and in the Gaining of Selective Places for Scottish Children and those in England and Wales 150
Duncan, K. D., and Foster, R.: An Investigation of the Craig Reader Course .. 327
Elley, W. B.: The Role of Errors in Learning with Feedback 296
Eyre, J. H.: The Prediction of Vocational Suitability from Secondary Modern School Record Cards 48
Eysenck, S. B. G., Syed, I. A., and Eysenck, H. J.: Desirability Response Set in Children 87
France, N., and Wiseman, S.: An Educational Guidance Programme for the Primary School 210
Fisher, B.: The Social and Emotional Adjustment of Children with Impaired Hearing Attending Ordinary Classes 319

Freyberg, P. S.: The Efficacy of the Coloured Progressive Matrices as a Group Test with Young Children .. 171
Hallworth, H. J. and Waite, G.: A Comparative Study of Value Judgments of Adolescents .. 202
Hitchman, P. J.: The Validity and Reliability of Tests of Spoken English .. 15
Holloway, P. J.: The Effect of Lecture Time on Learning .. 255
Irvine, S. H.: Towards a Rationale for Testing Attainments and Abilities in Africa 24
Kline, P.: Extraversion, Neuroticism and Academic Performance among Ghanaian University Students .. 92
Lee, T.: A Null Relationship between Ecology and Adult Education .. 100
Ley, P., Spelman, M. S., Davies, A. D. M., and Riley, S.: The Relationships between Intelligence, Anxiety, Neuroticism and Extraversion .. 185
Lunzer, E. A.: Concept Growth and the Education of the Child, by J. G. Wallace .. 124
Maguire, U.: The Effects of Anxiety on Learning, Task Performance and Level of Aspiration in Secondary Modern School Children .. 109
Marsh, J. F.: An Investigation into the Differences Between Children in the Formation of Concepts of Town and Country .. 333
McIntyre, D., Morrison, A., and Sutherland, J.: Social and Educational Variables Relating to Teachers' Assessments of Primary School Pupils .. 272
Mitchell, S., and Shepherd, M.: A Comparative Study of Children's Behaviour at Home and at School .. 248
Morrison, A., and Hallworth, H. J.: The Perception of Peer Personality by Adolescent Girls .. 241
Mundy, J. M.: Junior Children's Responses to the Murray Thematic Apperception Test and the Cattell Personality Questionnaire .. 103
Murdoch, W. F.: The Effect of Transfer on the Level of Children's Adjustment to School .. 332
Musgrove, F.: The Social Needs and Satisfactions of Some Young People at Home, Youth Clubs, and at Work, Part I .. 61
Musgrove, F.: The Social Needs and Satisfactions of Some Young People: Part II At School .. 137
Ojha, A. B. Kelvin, R. P., and Lucas, C. J.: A Note on Season of Birth and Intelligence .. 94
Peel, E. A.: A Study of Differences in the Judgments of Adolescent Pupils .. 77
Pollock, G. J.: The Opinions of Science Teachers on the Objectives of Teaching Science .. 112
Poppleton, P. K., and Brown, P. E.: The Secular Trend in Puberty: Has Stability been Achieved? .. 95

Radford, J.: Verbalisation Effects in a 'Non-Verbal' Intelligence Test .. 33
Reed, G. F.: The Association of Auditory High Frequency Weakness with Verbal and Written Comprehension and Expression .. 118
Riley, C. S.: The Relationship Between Reading Ability and Verbal Intelligence Test Performance .. 117
Rose, J. I.: A Sociometric Comparison of Well-Behaved and Badly-Behaved Classes 331
Rushton, J.: The Relationship Between Personality Characteristics and Scholastic Success in Eleven-year-old Children .. 178
Savage, R. D. Personality Factors and Academic Attainment in Junior School Children .. 91
Savage, R. D., and O'Connor, D. J.: The Assessment of Reading and Arithmetic Retardation in the School .. 317
Sigel, I. E., Roeper, A., and Hooper, F. H.: A Training Procedure for Acquisition of Piaget's Conservation of Quantity: A Pilot Study and its Replication .. 301
Smith, F. V.: Contrary Imaginations, by Liam Hudson .. 334
Start, K. B.: The Relation of Teaching Ability to Measures of Personality .. 158
Taylor, P. H.. A Study of the Effects of Instructions in a Multiple-Choice Mathematics Test .. 1
Thyne, J M.: What is the Weight of an Examination Paper? .. 321
Yates, A. J The Relationship Between Level and Speed on Two Intelligence Tests 166
Yates, A. J.: Level, Speed and Personality Factors in the Intelligence Performance of Young Children .. 312

NOTES TO CONTRIBUTORS

1.—The Editor is glad to consider articles in the fields of educational and child psychology, whether written from the theoretical, clinical, experimental or psychometric standpoints. It is, however, desirable that statistical matter (formulae, tables, discussions of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers. Such concepts as the mean, percentiles, standard deviation, correlation, chi-squared, standard error, critical ratio, significance and reliability, and general or group factors, can be employed in the text. But more complex concepts and methods should be explained or confined to footnotes or an appendix, or published elsewhere.

Articles which have been, or are to be, published elsewhere cannot be accepted.

2.—*Length.* The usual range of length is from 2,500 to 4,500 words, and only very exceptionally will articles of more than 6,000 words be published. Short research reports of 500—1,000 words, such as outlines of theses, are welcomed. These should normally be submitted through the Head of the Department in which the research was carried out.

3.—*Arrangement.* Articles should be ***preceded*** by a summary or abstract of between 100 and 300 words, embodying the main conclusions. This should be typed directly after the title and author's name and place of work. A contents analysis, or list of headings, need ***not*** be included, but successive sections of the article should receive numbered headings, e.g. :

I. Introduction.
II. Outline of Previous Research.
III. Plan of the Investigation.
  (i) *Subjects and materials.*
  (ii) *Experimental Procedure.*
IV. Statement of Results.
V. Discussion and Conclusions.
VI. References.

4.—*Typing.* All papers must be typed (double or 1½-spaced) on one side of quarto or foolscap paper· There should be margins of at least one inch all round. Authors should consult articles in current numbers of the *Journal*, and copy their conventions in respect of capitalising the title and main section headings, italicising author's place of work and sub-section headings, indentation of paragraphs, the form of references and tables, etc.

5.—*Footnotes* should be few in number, and should be indicated in the text by using the signs *, †, ‡, §.

6.—*References* should normally be indicated by date following the author's name in the text, e.g., Spearman (1904), or (Spearman, 1904, p. 160), or—where there are two or more publications in the same year—Burt (1939a). However, it is occasionally more convenient to number the references throughout, and to type numbers in the text above the line, e.g., Spearman.[4] When there are two or three references only, these may be placed in footnotes in the text, instead of in a separate section at the end.

The list of references at the end should always be in alphabetical order of author's names, and—for any one author—by order of date. The following standard form of references should be observed :

Dearborn, W. F., and Rothney, J. W. M. (1941). *Predicting the Child's Development.* Cambridge, Mass. : Sci.-Art.

(Note that Ltd., etc., after publisher's names, also the numbers of pages in the book, can be omitted.)
Burt, C. (1939a). The relations of educational abilities. *Brit. J. Educ. Psychol.*, 9, 55-71.

(Note the absence of inverted commas and of capital initials in the titles of articles. The number of the volume is in arabic, not roman numerals. The part of the journal is not given, but page numbers are included in full.)

The World List of abbreviations (as given in the December issues of *Psychological Abstracts*) is used for journals.

7.—*Tables and figures* are costly to print, and should be used only when they save space and/or are essential for effective presentation. The same material should not be presented in a table ***and*** a graph. Each table or figure should be numbered (in arabic figures), on a separate sheet, and its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy, unruled paper, the lettering on graphs being in pencil. Lines in tables should be ruled in pencil. Each table or graph should have a short title, sufficient for understanding without searching in the text.

8.—*Corrections.* The Editor cannot consider for publication articles which are seriously deficient in presentation. The typescript submitted (after preliminary consultation, if desired) should represent the final form in which the author wishes the paper to appear. Since changes in proof are costly, out of all proportion to the original setting, authors must be charged for any changes, insertions or deletions other than printer's errors).

9.—*Reprints.* Authors receive fifty copies of their papers free. Extra copies, in multiples of fifty, may be had at cost price if the order is given when the proof is returned.

THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY ARTHUR SUMMERFIELD

CONTENTS OF VOLUME 57, PARTS 1–5, 1966

SATINDER, K. PAUL. Effects of intermodal stimulation of figural after-effects.

KINSBOURNE, M., WARRINGTON, ELIZABETH K. and JAMES, MERLE. Uncertainty and transitional probability in the span of apprehension.

PRITCHATT, D. and HOLDING, D. H. Guiding Deutsch's model in reverse.

MATTHEWS, W. A. Sutherland's two-stage theory of discrimination learning and its experimental support.

DALE, H. C. A. Positive and negative recency in two-choice guessing.

WOODHEAD, MURIEL M. Varying the number of alternatives in short-term recall.

ANTHONY, W. S. Working backward and working forward in problem-solving.

WETHERICK, N. E. The inferential basis of concept attainment.

TAJFEL, HENRI and BRUNER, JEROME S. The relation between breadth of category and decision time.

DAWES, ROBYN MASON. Memory and distortion of meaningful written material.

DAS, J. P. Relation between semantic satiation and verbal conditioning.

KENNEDY, R. A. and KEENE, A. G. The effect of short periods of food-deprivation on human performance.

EYSENCK, HANS J. and THOMPSON, WARREN. The effects of distraction on pursuit rotor learning, performance and reminiscence.

EYSENCK, H. J. and WILLETT, R. A. The effect of drive on performance and reminiscence in a complex tracing task.

CLARK, A. D. B. and COOPER, G. M. Age and perceptal-motor transfer in imbeciles: task complexity as a variable.

CLARKE, ANN M., COOPER, G. M and HENNEY, A. S. Width of transfer and task complexity in the conceptual learning of imbeciles.

HASAN, PARWEEN and BUTCHER, H. J. Creativity and intelligence: a partial replication with Scottish children of Getzels' and Jackson's study.

BURT, CYRIL. The genetic determination of differences in intelligence: a study of monozygotic twins reared together and apart.

BELL, C. R. and WATTS, ANNE N. Personality and judgments of temporal intervals.

BOE, ERLING E. and KOGAN, WILLIAM S. Social desirability in individual performance on thirteen MMPI scales.

HEIM, A. W. and WATTS, K. P. The Brook reaction test of interests.

MAIR, J. M. M. Prediction of grid scores.

JAHODA, GUSTAV. Geometric illustrations and environment: a study in Ghana.

PASTORE, NICHOLAS and TERWILLIGER, MARLENE. Induction of stereoscopic depth effects.

LOCKS, ERWIN. The contradiction of epiphenomenalism.

PUBLICATIONS REVIEWED.

Subscription 90s. net per volume of 4 parts
Orders may be placed with a bookseller or sent direct to the publishers

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London, N.W.1

AN INTRODUCTION TO EDUCATIONAL PSYCHOLOGY

E. Stones

A book for students of education presenting an integrated view of educational psychology which embraces the most recent findings in the field as well as the major aspects of previous work.

36s ; University Paperback 18s

PLEASE, MISS, CAN I PLAY GOD?

Joan Haggerty

'A direct and enthusiastic account of dramatic work with children in a Junior School in the East End of London . . . Written lightly, directly from experience, and with the kind of honesty one would wish all teachers could bring to their work.' *The Guardian.* 21s

---

MODERN TEACHING

A new series for training colleges and teachers

Edited by P. W. Musgrave

The first title in the series:

CHILD GUIDANCE AND THE SCHOOL

Isabella C. Maclean

This book is designed to increase understanding of the problems faced in school by both normal and abnormal children. The author discusses school problems and specific types of difficult children, making valuable suggestions of ways in which school and home co-operation can be effective in the prevention of the difficulties so often encountered by psychologists, parents and teachers. 18s ; Paperback 9s 6d

---

The Library of Educational Thought

GEORG KERSCHENSTEINER

His thought and its relevance today

Diane Simons

An informed account of the modern German system of vocational education, pioneered by Georg Kerschensteiner. The success of his system and its relevance to British industrial problems and to the implementing of the Industrial Training Act, make this book essential reading for everyone concerned with education in this country. 25s ; Paperback 12s 6d

METHUEN

STUDIES OF TROUBLESOME CHILDREN

D. H. STOTT

"The empirical findings offered in this book cannot be disregarded. They constitute a challenge to much current thinking and are of importance to all who are concerned with maladjusted behaviour in any of its forms."

The London Teacher 32*s net*

ABSENT

School Refusal as an Expression of Disturbed Family Relationships

MAX B. CLYNE

In this study, which is illustrated by case material from his own practice, a family doctor attempts to clarify the concept of school refusal, to establish differential diagnostic criteria, and to elucidate the dynamics and aetiology of the condition. 38*s net*

PSYCHOTHERAPY AND CHILD DEVELOPMENT

An Introduction for Students

JEAN BIGGAR

A child psychiatrist discusses, with a wealth of clinical examples, the formative part played by the experiences of the early years in establishing the foundations of character and of mental health. 25*s net*

IDENTIFICATION AND CHILD REARING

ROBERT R. SEARS, LUCY RAU & RICHARD ALPERT

"This is a thoroughly scientific psychological study. It is well written in a pleasant easy style. The whole planning of the book is excellent and beyond reproach. Though the study is purely psychological, the terms used are properly explained, and the basis of the theories mentioned is well discussed. Each chapter concludes with a good discussion and summary, and there is a good final chapter of conclusions . . .

This book can be heartily recommended to all interested in the psychological study of the basis of child behaviour, and of the methods by which such a study can be carried out." R. S. Illingworth, *British Medical Journal* 63*s net*

SCHOOLS FOR YOUNG OFFENDERS

GORDON ROSE

A comprehensive account of the development and structure of the approved school system in England and Wales, with some discussion of its methods and problems. The information presented is of particular value at a time when radical changes in this form of training are envisaged, in accordance with current government proposals concerning the treatment of young offenders.

International Library of Criminology No. 17 52*s net*

TAVISTOCK PUBLICATIONS

New 5th Edition

Munn's Psychology

NORMAN L. MUNN

Professor of Psychology at the University of Adelaide

In this, the fifth edition of Munn's standard work, extensive reorganization, compression of material which time has shown to be less significant, and deletion of early findings which have not been substantiated by recent research have made way for the inclusion of a wealth of new material within the same number of pages and in fewer chapters. The book presents an outstanding survey of modern scientific psychology which is perfectly suited to the introductory course in general psychology. It not only gives the student an understanding of psychological processes and their application, but also places the study of mental life in its proper relation to study in such fields as biology and social behaviour. The book contains 12 pages of colour plates, and over 450 photographs, charts, and line diagrams in the text.

5th edition (1966) 56*s*.

Overview of General Psychology

A Basic Programme

L. DODGE FERNALD, JR.
Bowdoin College, U.S.A.

PETER S. FERNALD
Springfield College, U.S.A.

This is a manual for students, written to accompany the fifth edition of Munn's *Psychology*, but it may be used with any introductory psychology text. It is a completely programmed review of the basic topics of introductory psychology. Over 1,600 programmed units for self-instruction are provided, followed by a diagnostic summary which enables the student to check his retention of the material.

Late 1966, 21*s*.

The Psychology of the Use of Audio-Visual Aids in Primary Education

G. MIALARET

Director of the Laboratoire de Psychopédagogie, University of Caen

The author, a leading exponent of audio-visual methods, gives a history of their uses and discusses the subject in its relationship to linguistics and semantics.

Published in association with UNESCO
(1966) 30*s*.

Experiments and Studies in General Psychology

L. DODGE FERNALD, JR.
Bowdoin College, U.S.A.

Ideally suited to the needs of first and second-year psychology students as well as those taking courses in education of various kinds, this ingenious manual presents a series of diverse and interesting explorations in the major areas of introductory psychology which do not demand special expensive laboratory equipment. The exercises are an integral part of the text and they compel the reader to work and think actively throughout. (1966) 15*s*.

Instructor's Manual (with answers) 7*s*. 6*d*.

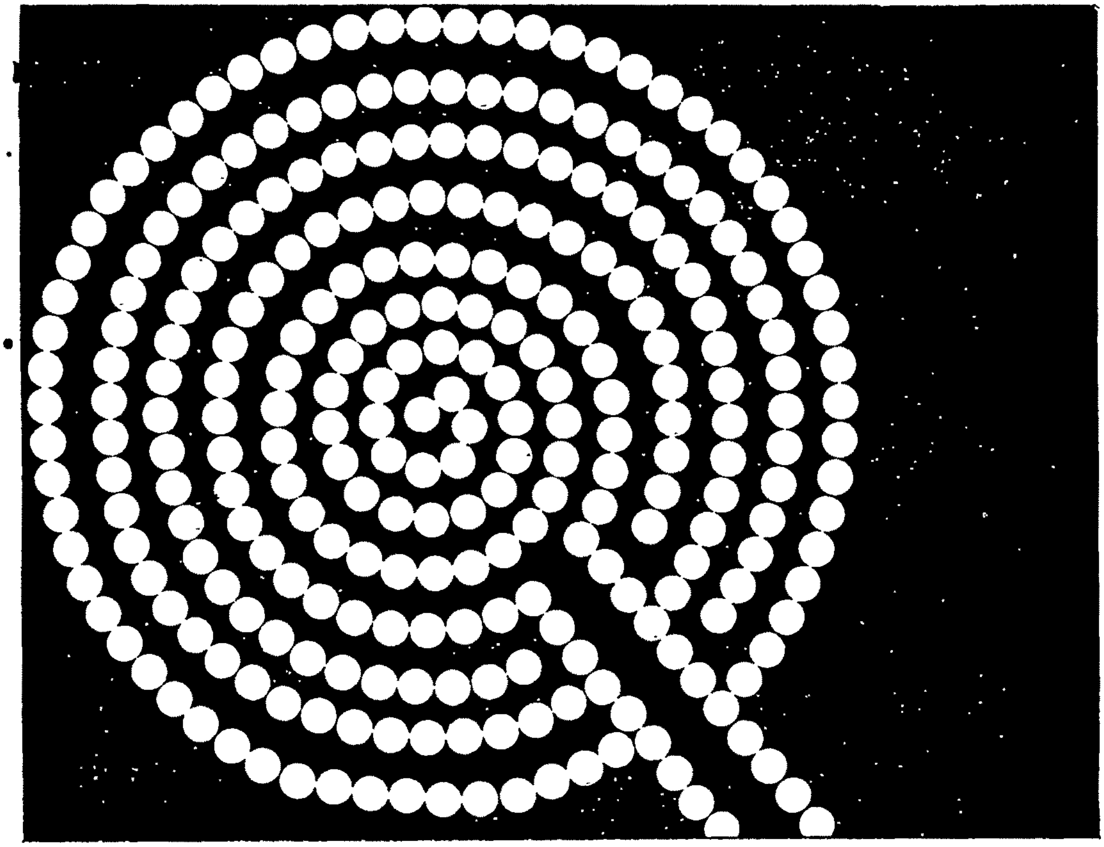
Porteus Maze Test

Fifty Years Application

STANLEY D. PORTEUS

This book, which marks the culmination of half a century of productive effort, is not a history but a discussion of the many problems to whose understanding the maze has contributed. (1966) 55*s*.

HARRAP
182 High Holborn, London, W.C.1



Announcing a major publishing event:

the appearance of the first four volumes of

PENGUIN MODERN PSYCHOLOGY READINGS

Attitudes Edited by Marie Jahoda and Neil Warren
Experiments in Visual Perception Edited by M.D. Vernon
Motivation Edited by Dalbir Bindra and Jane Stewart
Personality Assessment Edited by Boris Semeonoff

– Essential readings, from classical to contemporary work over the whole range of psychological thought and experiment, selected and edited by leading authorities.
– Each volume has a general and part introduction, further readings and a full index.
– At the low price of 8s 6d each they are within easy reach of every student's pocket.
– Over 25 further volumes are in active preparation under the general editorship of Professor B.M. Foss, Professor of Educational Psychology, London University.

Published by Penguin Books Ltd, Harmondsworth, Middlesex.

Write for a detailed catalogue or look at the books themselves in your bookshop now.

OCCUPATIONAL PSYCHOLOGY

Editor : ALEC RODGER

JANUARY and APRIL, 1966 VOLUME 40, Nos. 1 and 2

CONTENTS

Preception of Managerial Needs and Skills in Two National Samples. By FRANK A. HELLER and LYMAN W. PORTER.
Validating a Workshop Reporting Procedure. By F. A. SNEATH, G. C. WHITE and G. A. RANDELL.
A Study in Attitude Change. By A. J. M. SYKES.
Sick Leave Among Railwaymen Threatened by Redundancy: A Pilot Study. By A. CONSTANCE OWENS
The Selection of Maintenance Engineers for Data Processing Equipment. By A. P. WILLIAMS
Teaching Paired Associates: The Problem of Age. By EUNICE BELBIN and SYLVIA DOWNS
Personality Needs and Expected Benefits from A Management Development Program. By LEOPOLD W. GRUENFELD
Batch Size, Cycle Time and Setting Time as Determinants of Productivity in Skilled Machining Work. By A. B. HILL and J. M. B. THICKETT
Recognition of a Change in a Vibration Frequency, Felt Through the Fingers, by Naïve and Experienced Subjects. By E. N. CORLETT
Organisational Psychology. By GODFREY GARDNER
Book Reviews

Annual Subscription 40 *shillings, overseas* 50 *shillings*

National Institute of Industrial Psychology
14, Welbeck Street, London, W.1.

The British Journal of
Social and Clinical Psychology

Edited by MICHAEL ARGYLE and R. R. HETHERINGTON

Contents of Vol. 5, Part 3, September, 1966

WILLIAMS, M. The effect of Context on Schizophrenic Speech
SEALY, A. P. and CATTELL, R. B. Adolescent Personality Trends in Primary Factors measured on the 16 PF and the HSPQ Questionnaires through ages 11 to 13
FOULDS, G. A. 'Psychic : Somatic' Symptoms and Hostility
GIBSON, H. B. The validation of a technique for measuring delinquent association by means of vocabulary
CHILD, D. Personality and Social Status
HAAS, K. Social Correlates of Hostility in a College Sample
SPELMAN, M. S. and LEY, P. Knowledge of Lung Cancer and Smoking Habits
BRUCKMAN, I. R. The relationship between achievement motivation and sex, age, social class, school stream and intelligence
JOURARD, S. M. An exploratory study of body-accessibility
Book Reviews

The subscription price per volume of four parts is 90s. net. Orders may be sent to your bookseller or direct to the publishers

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200, Euston Road, London, N.W.1

AUSTRALIAN JOURNAL OF PSYCHOLOGY

Editor : I. K. WATERHOUSE, Macquarie University, N.S.W.

Vol. 18 *No.* 2 *August*, 1966

Yates, A. J. Data-processing Levels and Thought Disorder in Schizophrenia.

Provins, K. A. Environmental Heat, Body Temperature and Behaviour.

Adam, June. The Relationship between Visual Illusions and Figural After-effects.

Yates, A. J. and MacDonald, R. P. Reminiscence as a Function of Perceptual Search.

Landauer, A. A. A Note on the Effect of Instructions on Size Judgments with the Tunnel Illusion.

John, I. D. Intensity of Non-Key Stimuli in the Donders' c-Type Reaction.

Radcliffe, J. A. A Note on Questionnaire Faking with 16PFQ and MPI.

Wilson, P. R. and Patterson, J. R., and Lysons, A. M. Sex, Birth Order and Volunteering Behaviour.

Book Reviews

The Journal is published three times a year. Subscription rates 36s. (sterling), and $5·50 in U.S.A. Members of B.P.S. and A.P.A. should enquire about special subscription rates from the Secretary of their Society. Non-members should send enquiries to Melbourne University Press, Parkville, N.2., Victoria, Australia, or to its agents.

in U.K.
Cambridge University Press,
200 Euston Road,
LONDON, N.W.1

in U.S.A.
Cambridge University Press,
32 East 57th Street,
NEW YORK 10022, N.Y.

Enquiries from the Orient should be sent directly to Melbourne University Press.

THE BRITISH JOURNAL OF MEDICAL PSYCHOLOGY

Edited by Dr. THOMAS FREEMAN

Volume 39, *Part* 3. CONTENTS *September*, 1966

A. H. Crisp. Transference, Symptom Emergence, and Social Repercussion in Behaviour Therapy.

Leon Salzman. Memory and Psychoanalysis.

Peter Lomas. Ritualistic Elements in the Management of Child-birth.

D. Banister and Phillida Salmon. Schizophrenic Thought Disorder : Specific or Diffuse.

Andrew McGhie. The Subjective Assessment of Sleep Patterns in Psychiatric Patients.

G. Teruel. Considerations for a Diagnosis in Marital Psychotherapy.

L. Chertok. A Study of Tensions Among Psychotherapists.

Cawte, Nari Djagamara and M. G. Barrett. The Meaning of Subincision of the Urethra to Aboriginal Australians.

Reviews :

Subscription £4 10s. per volume of 4 parts. Single parts 30s. plus postage

Orders may be placed with any bookseller, or direct to :

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200, Euston Road, London, N.W.1

C.U.

The Birmingham Printers Ltd

PRINTERS AND BOOKBINDERS

High Class
Technical Publications

42-44 HILL STREET, BIRMINGHAM 5

Telephone: MIDLAND 0196-7

JOHN MACNAMARA, CM 35s.

Bilingualism and Primary Education

a study of Irish experience

This controversial study, based on an exhaustive examination of bilingual instruction on Irish primary school children, has international significance and makes this a book of the first importance for psychologists, psycholinguists, politicians and educationalists.

A technical story skilfully told by a dedicated social scientist. W. E. Lambert, Professor of Psychology, McGill University, in his introduction to the book.
'. . . probably the most valuable research into primary education in this country since the foundation of the State. A painstakingly accurate and unbiased educational study.' *The Irish Times*

EDINBURGH
UNIVERSITY PRESS
George Square, Edinburgh 8